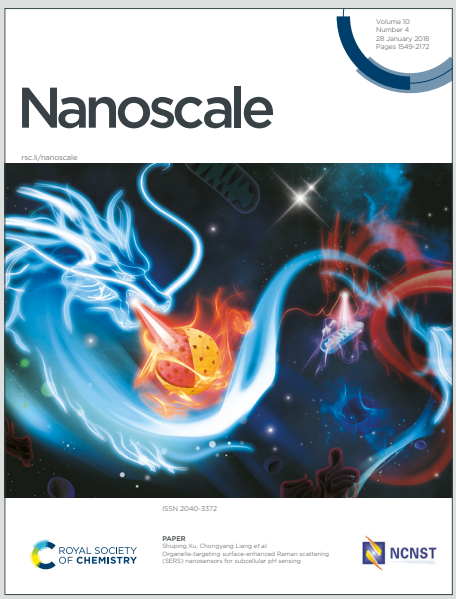
Nanoscale

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: N. BONO, B. Coloma Smith, F. Moreschi, A. Redaelli, A. Gautieri and G. Candiani, *Nanoscale*, 2021, DOI: 10.1039/D0NR09052B.



ROYAL SOCIETY OF CHEMISTRY

rsc.li/nanoscale

View Article Online
DOI: 10.1039/D0NR09052B

ARTICLE

In silico prediction of the *in vitro* behavior of polymeric gene delivery vectors

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Nina Bono[a], Bárbara Coloma Smith[a,b], Francesca Moreschi[a,b], Alberto Redaelli[b], Alfonso Gautieri[b,*] and Gabriele Candiani[a,*]

Non-viral gene delivery vectors has increasingly come under the spotlight, but their performace is still far from being satisfactory. Therefore, there is an urgent need for forecasting tools and screening methods to enable the development of ever more effective transfectants. Here, coarse-grained (CG) models of gold standard transfectant poly(ethylene imine)s (PEIs) have been profitably used to investigate and highlight the effect of experimentally-relevant parameters, namely molecular weight (2-10 kDa) and topologies (linear vs. branched), protonation state, and ammine-to-phosphate ratios (N/Ps), on the complexation and the gene silencing efficiency of siRNA molecules. The results from the *in vitro* screening of cationic polymers and conditions were used to validate the *in silico* platform that we have developed, such that the hits which came out of the CG models were of high practical relevance. We show that our *in silico* platform enables to foresee the most suitable conditions for the complexation of relevant siRNA-polycation assemblies, thereby providing a reliable predictive tool to test bench transfectants *in silico*, and foster the design and development of gene delivery vectors.

Introduction

Gene delivery by non-viral means has increasingly come under the spotlight as a major breakthrough in molecular medicine due to the development of therapeutics and preventatives,[1,2] such as Pfizer/BioNTech's and Moderna's COVID-19 vaccines,[3,4] and in basic research for the investigation of cellular mechanisms or the production of specific proteins.[5] The transfection process encompasses the delivery of foreign nucleic acids (NAs) into eukaryotic cells and the deliberate modulation of gene expression patterns in the form of up/downregulation or even silencing of a specific gene(s) of interest. Because anionic DNA and RNA cannot spontaneously cross the negatively charged cell membrane, the delivery of naked NAs is the simplest but rather ineffective way to transfer the genetic cargo to cells.[6] In this context, search in various scientific databases reveals that most publications in genetics, biochemistry, molecular and developmental biology relied on the use of cationic lipid or polymer transfectants.[7–10] These molecules are considered relatively safe, they display easily tunable physico-chemical properties, can be smoothly produced in large quantities, with high reproducibility and affordable costs, and exhibit an unlimited NAs ferrying capacity.[11–13] The preparation of non-viral gene delivery vectors is likely a simple yet effective process: cationic lipids or polymers spontaneously assemble with NAs to give rise to polyelectrolyte complexes called lipoplexes and polyplexes, respectively. Their size spans from few tens to several hundreds of nm in hydrodynamic diameter (D_H), such that they are internalized by target cells, and deliver their payload intracellularly.[13] Unfortunately, efficacy (i.e., transfection efficiency) and adverse effects (i.e., cytotoxicity) are intimately related,[13,14] such that there is no transfectant that is effective and non-toxic at once. The dilemma we face is best exemplified by considering the gold standard polymeric gene delivery vector poly(ethylene imine) (PEI).[15,16] The performance of PEI-based polyplexes depends on the molecular weight (M_W) and topology (i.e., linear or branched) of the polymer itself, and its degree of protonation (which is a function of the pH of the dispersant).[16,17] Yet, other features, such as the size and the surface charge of polyplexes, which are also a function of the complexation conditions, affect to some extent the *in vitro* and *in vivo* toxicity and efficacy of polyplexes. [18,19]

In this context, despite a large variety of gene delivery vectors have been designed, patented and eventually commercialized,[20,21] the development of more and more effective transfectants is a too lengthy, costly, and laborious process.[22] *In silico* approaches may represent a mean of high-throughput test bench of different transfectants and conditions before significant investment is made. Viewed in perspective, these tools may allow scientists to foretell the behavior of a transfectant before it is synthesized, combined in polyplexes with NAs, and tested for efficacy.[20] Although the size of polyplexes and the time-scale for their formation have been posing a major computational hurdle, the systems addressable by computational modeling are nowadays increasingly complex. Nonetheless, the use of classical atom-level (AA)

[a] GenT LAB, Dept. of Chemistry, Materials and Chemical Engineering "G. Natta", Politecnico di Milano, 20131 Milan (Italy). E-mail: gabriele.candiani@polimi.it
[b] Biomolecular Engineering Lab, Dept. of Electronics, Information and Bioengineering, Politecnico di Milano, 20131 Milan (Italy). E-mail: alfonso.gautieri@polimi.it

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

molecular dynamics (MD) approaches to provide detailed mechanistic insights on the complexation mechanism[23–29] is restricted to polyplexes made of one or very few NA molecules[30–35] complexed with very low M_W (very LM_W, less than 2 kDa) transfectants only.[36,37,17,28,38,39] Conversely, coarse-grained (CG) simulations are well-suited to model polyplexes of experimentally-relevant size and to simulate different transfectants and conditions (e.g., M_W, branching, charge, salt concentration).In this context, different CG models of LM_W PEI have recently been released.[36,40–42]

In this work we aimed to fill the gap in the current understanding of the performances of gene delivery vectors, by implementing an *in silico* platform able to predict and identify the most effective transfectant(s) and complexation conditions in terms of ammine-to-phosphate ratios (N/Ps), i.e., the moles of the transfectant used to complex NAs. To this purpose, in this proof-of-concept study we have identified and analyzed systems that are experimentally relevant and easily treatable *in silico*. They consist of polyplexes made of siRNA and PEIs in the most experimentally relevant M_W and molecular topologies (i.e., linear and branched, *l*PEI and *b*PEI). PEI is the gold standard polymeric transfectant.[15,16] By the same token, siRNA has been widely used (especially in combination with PEI) over the last two decades to silence the expression of target genes *in vitro* and *in vivo*.[43] Of note, siRNA molecules can be easily handled computationally.

We herein carried out complexation and transfection assays and pinpointed *in vitro* the most effective transfectants and conditions. Alongside, siRNA-PEI complexations were run *in silico* in the exact same conditions as *in vitro*. The hits that came out of the *in silico* screening allowed us to come up with a predictive platform to test bench *in silico*, in a reasonable time-frame, a variety of transfectants and conditions, and expedite the development of ever more effective gene delivery vectors.

2. Materials and Methods

2.1 *In vitro* materials

*b*PEIs with M_W of 2 kDa (cat. no. 06089-100) and 10 kDa (cat. no. 19850) and 2.5 kDa *l*PEI (cat. no. 24313) were from Polysciences (Eppelheim, Germany), while 10 kDa *l*PEI was from Merck Life Science (Rome, Italy). Lyophilized double-stranded (ds) anti-GFP siRNA (siGFP; 5'-GGCUACGUCCAGGAGCGCACCdTdT-3'), ds siRNA scramble (5'-CUUACGCUGAGUACUUCGAdTdT-3') and 5× siMAX buffer (30 mM HEPES, 100 mM KCl, 1 mM $MgCl_2$, pH 7.3) were purchased from Eurofins Genomics (Ebersberg, Germany). GFP-expressing 293-cells (GFP-293 cells; cat. no. AKR-200) were from CellBiolabs Inc. (San Diego, CA, USA). AlamarBlue Cell Viability assay® was from Life Technologies (Monza, Italy), while BCA Protein Assay Kit was from ThermoFisher (Monza, Italy). All the other chemicals were from Merck Life Science, if not differently specified.

2.2 Preparation of siRNA solutions

Lyophilized siRNAs were first solubilized in 400 µL of 1× siMAX sterile buffer to reach a final concentration of 100 pmol µL^{-1}, according to manufactures' instructions. siRNA solutions were next diluted in sterile 1× siMAX to reach the final concentration of 100 pmol µL^{-1} (corresponding to 0.1 µg µL^{-1}), then the concentration was assessed by measuring the OD_{260} by means of a spectrophotometer (Nanodrop 2000c, Fisher Scientific, Illkirch, France), and stored at -20 °C until use.

2.3 Preparation of PEI solutions

PEI solutions were prepared in sterile water for injection (injection-H_2O) at a concentration of 1 mg mL^{-1}, and the pH was adjusted to 7.0. Afterwards, the solutions were diluted in injection-H_2O to give PEI stock solutions of 0.86 mg mL^{-1} (corresponding to an amine concentration ([N]) of 20 mM).[22] Stock solutions were stored at 4°C until use.

2.4 Preparation of polyplexes

Before complexation, PEI and siRNA solutions were warmed to 30 °C, then vortexed for 10 s at 800 rpm before use. Complexes were obtained dripping the siRNA to PEI solutions (at the stoichiometric ratio of 1:7 (v/v)) prepared at the desired polymer concentration in 150 mM NaCl (pH 7.4) to yield different N/Ps. N/P is defined as the amine moles (N, cationic moiety) of the polymers with respect to the phosphate moles (P, anionic moiety) of a given quantity of siRNA. Suspensions were next incubated for 20 min at 30 °C prior to use.

2.5 Evaluation of siRNA complexation: SYBR Green I exclusion assay

The ability of each cationic polymer to bind and complex siRNA was evaluated by means of a fluorescence-exclusion titration assay. Briefly, polyplexes were invariably prepared by mixing 0.20 µg of siRNA with 12 µL of PEI solutions at different concentrations, yielding different N/Ps (i.e., 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 5, 6, 10). Following 20 min-incubation at 30 °C, 2 µL of 200× SYBR Green I were added to polyplexes, then incubated for 15 min at 30 °C and kept in the dark. Afterwards, the suspensions were diluted 1:1.5 in 150 mM NaCl, then fluorescence measurements (λ_{ex} = 497 nm, λ_{em} = 520 nm; n = 3 per condition) were performed with a Synergy H1 spectrophotometer (BioTek, Italy) in 384-multiwell black plates. Free siRNA (%) expresses the amount of free (unbound) siRNA present in polyplexes prepared at varying N/Ps, and normalized with respect to the fluorescence of uncomplexed (naked) siRNA, according to the following equation:

$$in\ vitro \text{ free siRNA } (\%) = \left(\frac{F_{sample} - F_{blank}}{F_{naked\ siRNA} - F_{blank}}\right) \times 100$$

where *F* is the recorded fluorescence.

Complexation curves were fitted with a sigmoidal dose-response model.

2.6 Physico-chemical characterization of siRNA-PEI polyplexes

For physico-chemical characterization, 40 µL of siRNA-PEI complexes were prepared as described above. Briefly, 0.5 µg of siRNA were complexes with 35 µL of PEI solutions (prepared in 150 mM NaCl) to yield N/P 10. Afterwards, the suspensions were diluted in 150 mM NaCl to reach a final volume of 500 µL and allowed to equilibrate for 5 min before measurements. The size (expressed in terms of D_H), and the surface potential (expressed as zeta potential, ζ_P) of polyplexes were measured at 25 °C by means of Dynamic Light Scattering (DLS) and Electrophoretic Light Scattering (ELS) techniques, respectively, using a Malvern Zetasizer Nano instrument (Malvern, Italy) fitted with a 5 mV HeNe laser (λ = 633 nm) and a scattering angle of 173°. The polydispersity index (PDI) was used to indicate the size distribution within the sample.

2.7 *In vitro* cell transfections

2.7.1 Cell culture

Mycoplasma-free GFP-293 cells were routinely cultured and expanded in T75 flasks in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 1 mM sodium pyruvate, 10 mM Hepes buffer, 100 U mL^{-1} penicillin, 0.1 mg mL^{-1} streptomycin, 2 mM glutamine and 10% (v/v) fetal bovine serum (FBS) (hereafter referred to as complete DMEM, cDMEM) incubated in a humidified atmosphere and 5% CO_2 at 37 °C (hereafter referred to as standard culture conditions).

2.7.2 Transfection assays

The timeline of transfection experiments is reported in **Figure S1**. For transfection assay, GFP-293 cells were seeded onto 96-wells plates at a density of 5×10^4 cells cm^{-2} and cultured in standard culture conditions. Twenty-four hrs post-seeding (t_{0h}), 34 pmol/well (corresponding to 250 ng/well) of siGFP were complexed with PEIs to yield different N/Ps (i.e., 5, 10, 15 and 30), as described here above. Polyplexes were allowed to form for 20 min at 30°C, then incubated with cells for 4 hrs in a final volume of 100 µL/well of serum-free cDMEM. Following 4 hr-incubation in the presence of polyplexes (t_{4h}), the medium was replaced with 100 µL/well of fresh cDMEM, then cells were cultured for additional 44 hrs in standard culture conditions (t_{48h}). Experiments were run at least in triplicates. In parallel, GFP-293 cells (≥ 3 wells per conditions) transfected with polyplexes prepared at the same NPs but harboring the scramble siRNA sequence (i.e., a non-functional siRNA designed not to target any gene) were used as negative controls.

The cytotoxicity was evaluated by means of the Alamar Blue assay®, according to the manufacturer's instructions. Briefly, the cDMEM was replaced with 100 µL/well of 1× resazurin dye solution in cDMEM. Next, plates were incubated in standard culture conditions for 2 hrs and at t_{48h} the fluorescence was read by means of a Synergy H1 spectrophotometer (λ_{ex} = 540 nm; λ_{em} = 595 nm). The viability of non-transfected cells (CTRL) was set to 100%, and the cytotoxicity was calculated according to the following equation:

$$\text{cytotoxicity (\%)} = 100\% - \text{viability (\%)}$$

$$= \left(1 - \frac{F_{\text{transfected cells}}}{F_{\text{CTRL}}}\right) \times 100$$

Cell silencing was evaluated by measuring the GFP fluorescence signal in cell lysate. Briefly, at t_{48h}, cells were washed with PBS, lysed with 110 µL/well of Cell Culture Lysis Reagent (Promega, Milan, Italy), then plates were stored at -80°C overnight. The fluorescence of 20 µL of cell lysates was next measured by means of the Synergy H1 microplate reader (λ_{ex} = 488 nm; λ_{em} = 509 nm). *F* values were normalized to the total protein content of each sample, as determined by the BCA assay. *F*/mg of proteins for non-transfected cells (CTRL) was set to 100%, and the transfection efficiency (TE) was expressed according to the following equation:

$$\text{TE (\%)} = \left(1 - \frac{\text{F/mg of proteins}_{\text{transfected cells}}}{\text{F/mg of proteins}_{\text{CTRL}}}\right) \times 100$$

For each condition, cell silencing (%) was determined as follows:

$$\text{GFP silencing (\%)} = \text{TE(\%)}_{\text{siRNA antiGFP}} - \text{TE(\%)}_{\text{siRNA scramble}}$$

2.8 CG models of siRNA and PEIs

The ds-siRNA model was the siGFP with 2-dT 3'-overhangs: (GGCUACGUCCAGGAGCGCACC)TT. siRNA atomistic model was constructed in its canonical B form with a total charge of -46 *e* by using *Nucleic Acid Builder*.[44] siRNA CG model was generated from the atomistic structure using the Martini FF RNA extension.[45] In the molecular models, the overhangs were made for simplicity by a pair of uracil nucleobases instead of thymine. The CG models of PEIs relied on the recently published work by Mahajan *et al.*, which reported the parameters for 0.6 kDa *l*PEI and *b*PEI.[36] Each and every CG bead contained C-C-N atoms and the corresponding hydrogens. The beads were labeled *P*, *S* and *T*, depending on the presence of primary, secondary and tertiary amines, respectively (**Figure 1a**). In the case of charged beads, the letter *q* was appended to the bead label (i.e., *Sq* and *Pq* for protonated secondary or primary amines).

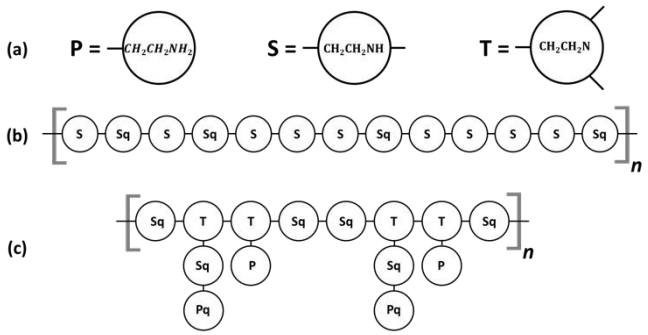


Figure 1. a) Coarse grained beads (and building blocks used to generate the CG models of **b)** *l*PEI and **c)** *b*PEI. Topologies and protonation were defined to match NMR and titration data. Each building block has a M_W of ~0.6 kDa. The resulting polymers considered in this work are: 2.5 kDa *l*PEI, 2 kDa *b*PEI, 10 kDa *l*PEI and 10 kDa *b*PEI.

We constructed CG molecular models of experimentally relevant PEIs. We built two different *l*PEI models of 2.5 kDa and 10 kDa by replicating the building block (M_w=0.6 kDa, **Figure 1b**)

DOI: 10.1039/D0NR09053B

by 4 and 16 times, respectively. Once obtained the desired M_W, both *l*PEI models were capped with a protonated terminal primary amine, resulting in CG models made of 53 (2.5 kDa *l*PEI) and 209 beads (10 kDa *l*PEI). The *l*PEI protonation at pH 7.4 was set at 32.5% according to titration data.[46]

The *b*PEI models were generated through the definition of a building block comprising 14 beads, and with a M_W of 0.6 kDa (**Figure 1c**). In order to define the degree of branching, the building block topology was designed according NMR data, that is, *b*PEI displayed 32% of primary amines, 40.6% secondary amines, and 27.4% tertiary amines. According to titration assays, the overall protonation at pH 7.4 was set to 55%, resulting from 55% and 100% of primary and secondary amines, respectively, which were protonated, while there were no tertiary amines that are protonated.[46] The *b*PEI models of 2 kDa and 10 kDa were generated by replicating the building block by 3 and 16 times, respectively. Once obtained the desired M_W, both *b*PEI models were added with a protonated terminal primary amine, resulting in models of 43 and 225 beads, respectively. Bonded interactions between CG beads (i.e., bonds, angles and dihedrals) are parameterized based on recently published parameters for PEI[36] (see also Electronic Supplementary Information (ESI)). Non-bonded interactions (electrostatic interactions and van der Waals interactions) are addressed by using MARTINI bead type Qd for positively charged beads, and bead type P2 for uncharged beads.[36]

2.9 Polyplex models

For each of the four PEI models considered, we generated 11 different systems with N/Ps ranging from 0.5 up to 10. In each system, the number of siRNA molecules was kept between 8 and 10, and the number of PEI molecules was chosen accordingly in order to obtain the desired N/P (**Table 1**).

Table 1. Summary of simulated systems

	2.5 kDa *l*PEI		2 kDa *b*PEI		10 kDa *l*PEI		10 kDa *b*PEI	
N/P	N_{siRNA}	N_{PEI}	N_{siRNA}	N_{PEI}	N_{siRNA}	N_{PEI}	N_{siRNA}	N_{PEI}
0.5	9	4	9	5	9	1	10	1
1	9	8	9	10	9	2	10	2
1.5	9	12	10	16	9	3	10	3
2	9	16	8	17	9	4	10	4
2.5	9	20	9	24	9	5	10	5
3	9	24	9	29	9	6	10	6
3.5	9	28	8	30	9	7	10	7
4	9	32	9	39	9	8	10	8
5	9	40	9	49	9	10	10	10
6	9	48	9	58	9	12	9	11
10	8	71	8	86	9	19	9	18

Every system was generated with molecules randomly distributed in a cubic simulation box with a side length of 20 nm. Systems were then solvated with MARTINI water beads[47] and counterions (Na^+ and Cl^-) were inserted to faithfully replicate the experimental conditions (150 mM NaCl). The final models comprised ≈ 70,000 beads.

2.10 MD simulations

The 44 molecular systems were simulated with GROMACS 2019 package[48] following protocols used in previous studies.[49–52] All the systems were minimized for 5,000 steps using the steep-descent method and equilibrated for 1 μs under constant pressure and temperature (NPT). In the NPT simulations, the temperature was kept constant at 300 K using a velocity rescaling thermostat with a time constant of 1.5 ps.[53] The pressure coupling was maintained at 1 bar, employing the Berendsen barostat[54] with time constant of 3 ps and compressibility $3 \cdot 10^{-4}$ bar^{-1}. A neighbor list was maintained within a radius of 1.2 nm using the Verlet scheme and was updated every 10 time-steps. Electrostatic interactions were calculated using the reaction field scheme with a cut off of 1.2 nm and a relative dielectric constant of 15, while van der Waals interactions were treated with the potential-shift scheme with a cut off at 1.2 nm. Each model was simulated for $50 \cdot 10^6$ steps with a time step of 5 fs, resulting in 250 ns of MD. Due to the scale factor of 4 in CG simulations,[55] the effective simulation time resulted to be 1 μs. The analyses of the simulations were performed using GROMACS tools and in-house tcl-scripts in VMD.[56]

2.11 Statistical analysis

Statistical analyses were carried out with GraphPad version 8 (GraphPad software, La Jolla, CA, USA). Comparisons between groups were performed with one-way analysis of variance (ANOVA) and multiple t-tests. Significance was retained when $p < 0.05$. Data are expressed as mean ± standard deviation (SD). Experiments were performed at least in triplicate.

3. Results and Discussion

This work was aimed at taking advantage of MD simulations as a predictive tool of the polyplex behavior.

In order to implement and validate a computational platform able to design the best conditions for siRNA-polycation complexation and eventually transfection, we first carried out *in vitro* characterization of polyplexes made of siRNA and the gold standard transfectant PEI, which was testes selected at the M_w and conditions that are the most relevant for gene delivery.[57] *In silico* parameters were chosen to closely match the experimental conditions.

3.1 siRNA complexation by means of PEIs

When dealing with the evaluation of cationic polymers for gene delivery, one shall determine experimentally the minimum ratio (N/P) at which NAs are fully complexed by a given polycation to give rise to effective assemblies, as no naked and ineffective

View Article Online
DOI: 10.1039/D0NR09052B

siRNA is delivered to cells.[22] Besides, because non-viral gene delivery particles are formed by electrostatic interactions,[58,59]

View Article Online
DOI: 10.1039/D0NR09052B

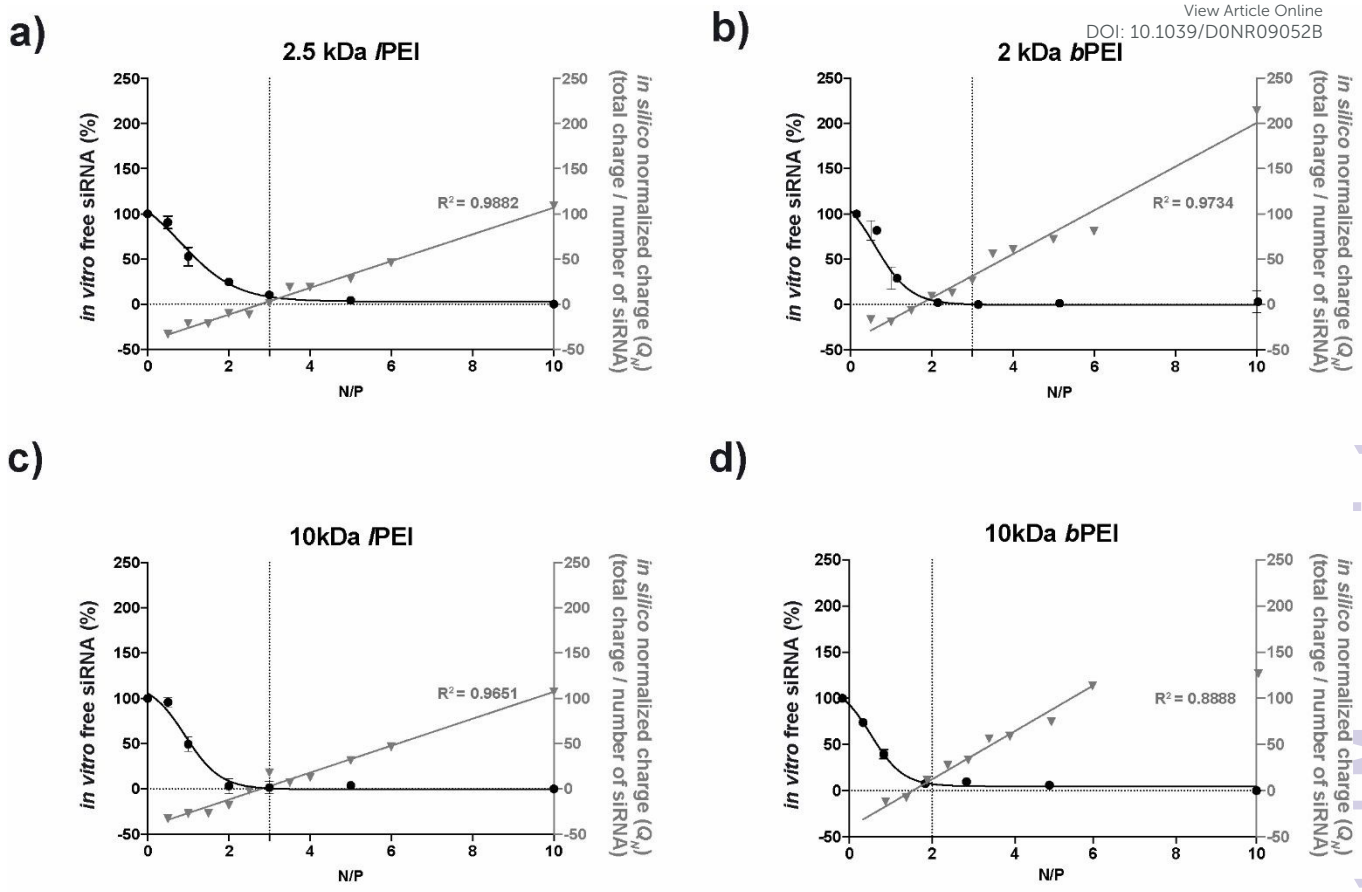


Figure 2. *In vitro* free siRNA (%) versus *in silico* normalized charge (Q_N) for each polyme over th *vitro* results refer to the black left Y axis, while in silico results refers to the grey right Y axis. *In vi* siRNA and fitted with a sigmoidal dose-response model.

they are sensitive to the composition of the medium (i.e., the saline composition, the ionic strength, and the pH) in which the complexation occurs. We decided to carry out all the experiments using one of the most widely used dispersing media, that is 150 mM NaCl at pH 7.4.[60–62]

The binding ability of PEIs towards siGFP was investigated by means of the SYBR Green I exclusion assay (**Figure 2**). Such a probe was used because it is known to exhibit a strong fluorescence enhancement upon binding to ds NAs, such as siRNA, thus providing high signal-to-noise ratios.[63]

In our experiments, 100% fluorescence was thus assigned to the free siRNA condition (N/P 0), where the maximum amount of NAs could be intercalated by the SYBR Green molecules. By increasing the N/P, i.e., by rising the number of polymer units added to the siRNA, less NAs were accessible to the probe and the fluorescence decreased as a consequence. Of note, each and every PEI had a high complexation capacity for siRNA.

Specifically, regardless of the M_w, *b*PEIs were slightly more effective in binding siRNA molecules than *l*PEIs, as they showed a minimum effective N/P for complete NAs complexation of 2, while the latter polymers did fully condense siGFP at N/P 3.

From a computational perspective, we analyzed the data extracted from the CG models and compared the results with experimental findings. Within the time-frame of the MD simulations we observed the shaping of siRNA-PEI complexes (see videos in ESI). Depending on the N/P and the M_w of the polymers, different siRNA-PEI polyplexes or clusters were apparent at the end of the 1 μs MD trajectory. To define the different clusters, we used a distance cut-off of 0.53 nm, as indicated in previous works.[42] In order to rationalize the data from MD simulations, we came up with a new index that made some genuine findings meaningful. We counted the number of polyplexes formed at the end of the MD simulations under any given condition tested, we selected the largest one and calculated its total charge as the algebraic sum of negative and positive charges which were found within the cluster itself. Anionic charges accounted for beads representing phosphate groups of the siRNA, while cationic charges took into account any protonated amine unit each polymer model (Qd beads). The overall charge of the cluster was then normalised with respect to the number of siRNA molecules belonging to it, according to the following equation:

$$\textit{in silico}\ \text{normalized charge}\ (Q_N) = \frac{[\text{total charge}]_{\text{largest polyplex}}}{[\text{number of siRNA}]_{\text{largest polyplex}}}$$

The analysis of the Normalized Charge (Q_N) as a function of the N/P (**Figure 2**) shows that the largest cluster displayed a net negative charge for small numbers of polymer units (i.e., low N/Ps), but Q_N increased almost linearly with N/P values, and it became largely positive for greater N/Ps. The dashed line marking the zero charge on the Q_N scale in **Figure 2** helps pinpointing the N/P values corresponding to the transition point, that is, the isoelectric condition at which the overall Q_N of the largest cluster was neutral for the four polycations. Altogether, **Figure 2** depicts the comparison between the Q_N of

the cluster, as obtained from MD simulations, and the complexation results as for the experimental SYBR Green I-exclusion assay. For any polymer topology and M_W tested, we matched the results of *in silico* predictions and experimental findings and found a definite correspondence in the determination of the minimum amount of polymer required to achieving effective NAs complexation, that is, the N/P at which polyplex neutrality was attained. As set forth herein above, the determination of this value is of very practical importance. On the other hand, the lowest N/P which allows to attain the greatest complexation of NAs seldom gives rise to the generation of the most effective gene delivery particles, in agreement with published data.[22]

Nevertheless, the overall concordance between *in vitro* and *in silico* findings discloses CG simulations as a novel and effective forecasting tool to explore the NA complexation ability of polymeric transfectants with different M_W and topologies.

3.2 *In vitro* cell transfection with siRNA-PEI polyplexes

Gene silencing experiments with PEIs were carried out to bring out the relevance of the cationic polymers and conditions explored in this manuscript and to highlight the most effective polymer(s) and condition(s) in transfection.

Of note, polyplexes made of 10 kDa PEIs and siGFP induced significant gene silencing (**Figure 3a**), while LM_w PEIs were unable to do so, as their transfection efficiencies was below the detection threshold (not shown).

The transfection efficiency and cytotoxicity of 10 kDa PEIs was dependent, to some extent, on the N/P.[16] Indeed, both kinds of siRNA-PEIs were barely toxic but rather ineffective in transfection at the lowest N/P tested (N/P 5). Instead, any type of polyplex attained its respective gene silencing maximum at N/P 10, and no significant increase in transfection efficiency and cytotoxicity was observed at higher N/Ps (**Figure 3**). Of note, similar cytotoxicity trends were found for any kind of control siRNA scramble-bearing polyplex (**Figure S1**). This implies that the cytotoxic effects of polyplexes were unrelated to the siRNA sequence delivered, but the toxicity was specific to the polymer used and N/P tested.

Nevertheless, irrespective of the N/P considered, the transfection efficiency of 10 kDa *b*PEI polyplexes was much greater than 10 kDa *l*PEI.

Overall, these results pointed to 10 kDa PEIs at N/P 10 as practically relevant for gene silencing purposes. On the other hand, this also means that just two out of four polymers that were selected are of experimental relevance, while the other half can profitably be used to benchmark larger M_w PEIs *in silico* and *in vitro*.

Nevertheless, it is worth noting that cell transfection is undoubtedly an example of extreme complexity. It encompasses several steps following polycation-NA complexation, that is, the interaction of polyplexes with the cell membrane and their internalization, and the intracellular release of NAs.[10,64] The size and time-scale of these phenomena naturally hamper their investigation through fully atomistic simulations. It is however generally known that the transfection behavior of polyplexes is dependent, at least in part, on their physico-chemical features (e.g., charge, size).[65,66]

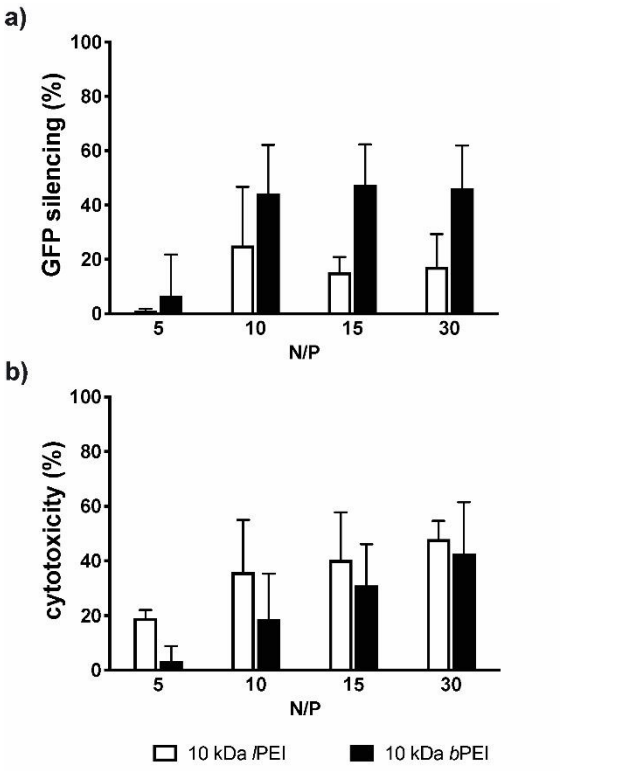


Figure 3. a) GFP silencing (%) and **b)** cytotoxicity (%) evaluated on GFP-293 cells, of siRNA-PEIs complexes (10 kDa *l*PEI = white bars; 10 kDa *b*PEI = black bars) prepared at different N/Ps (5, 10, 15, 30). Results are expressed as mean ± SD (n ≥ 3) ($, #: $p < 0.05$ *vs*. any other N/P).

3.3 Characterization of siRNA-PEI polyplexes

In order to draw a more comprehensive picture on polyplex features and behavior, siRNA-PEI complexes at N/P 10 in 150 mM NaCl were further characterized in terms of overall dimensions (D_H) and electr(okinet)ic potential (ζ_P).

As depicted in **Figure 4**, different polyplexes showed specific D_H and ζ_P as a function of the M_W and topology of the PEI used.

These results may explain the polymer-specific behavior in transfection of each kind of polyplex. As a rule of thumb, the greater the ζ_P (to some extent) and the more favourable the interactions between polyplexes and cells. By the same token, the larger D_H of polyplexes, the greater their settling over cells.[19] In line with this, 10 kDa *b*PEI-based polyplexes showed the most positive ζ_P (+18 ± 11 mV; $p < 0.05$ *vs.* LM_w counterparts) and the largest D_H (556 ± 100 nm; $p < 0.05$ *vs.* any other experimental group), and were the most effective in transfection, indeed. Instead, polyplexes made of 10 kDa *l*PEI had less positive ζ_P (+11 ± 3 mV) and were significantly smaller (412 ± 43 nm), such that their transfection efficiency was dramatically impaired.

Finally, we can speculate that both LM_W PEIs were ineffective in transfection because of the slightly negative charge (average ζ_P = -3 mV) and, at least in part, the relatively small dimension (average D_H = 368 nm) of the resulting polyplexes.

View Article Online
DOI: 10.1039/D0NR09052B

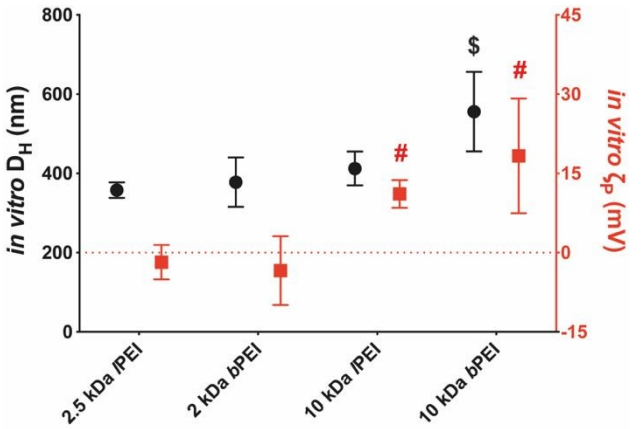


Figure 4. Size, expressed as hydrodynamic diameter (D_H (nm); black dots) and surface potential, expressed as zeta potential (ζ_P (mV); red squares) of siRNA-PEI polyplexes at N/P 10 (\$: $p < 0.05$ *vs.* any other experimental groups: #: $p < 0.05$ *vs.* LM_W PEIs).

Although one might be tempted to draw parallels between findings on Q_N and ζ_P, they depict features that are inherently different from each other. Indeed, the ζ_P, also termed as electrokinetic potential, is the potential at the slipping/shear plane of colloid particles moving under an electric field.

It should be noted that the ζ_P reflects the potential difference between the Electric Double Layer (EDL) of electrophoretically mobile particles and the layer of dispersant around them at the slipping plane. This means that ζ_P deals with the surface potential of polyplexes. Conversely, Q_N is an index of the overall charge of the largest polyplex (only) present in a model system of finite size, such as the simulation box. In light of the above, it is apparent that the two indexes are unrelated to each other. By the same token, but for different reasons, it would be improper to compare experimental and *in silico* measurements about the size of polyplexes. In fact, CG modelling allows to simulate a significant but small part of the real system (i.e., the whole polyplex suspension), which instead consists of a far greater number of siRNA and polymer molecules, actually. To a different extent, this drawback unfortunately applies to each and every *in silico* approach which deals with large systems.

In the same conditions as above (e.g., N/P 10), it is deemed reasonable that, the greater the D_H of polyplexes, the fewer their number and vice versa. Because DLS and other experimental technologies cannot provide an estimate of the number of polyplexes in the system, we decided once again to make full use of the potential of CG simulations to do so.

Figure 5 shows the number (panel a) and the appearance (panel b) of polyplexes at N/P 10 in the last 100 ns of MD simulations, as a function of the M_W and topology of PEI.

It is worth noting that LM_W PEIs gave rise to the formation of several polyplexes, typically composed of one or two siRNA oligos wrapped up with single PEI molecules. This effect was particularly evident for *b*PEI. Conversely, 10 kDa PEIs did form very few but considerably large aggregates which comprised early all siRNA molecules of the system. In the case of 10 kDa *b*PEI, the simulations showed the formation of one single large cluster. Again, this information is complementary, yet not less fundamental than that obtained experimentally on the size and PDI of the aggregates (**Figures 4** and **S3**). In this regard, we can speculate that the transfection behavior of a given polymer at a specific N/P may be inversely related to the number polyplexes that it is able to generate, as apparent from the CG simulations.

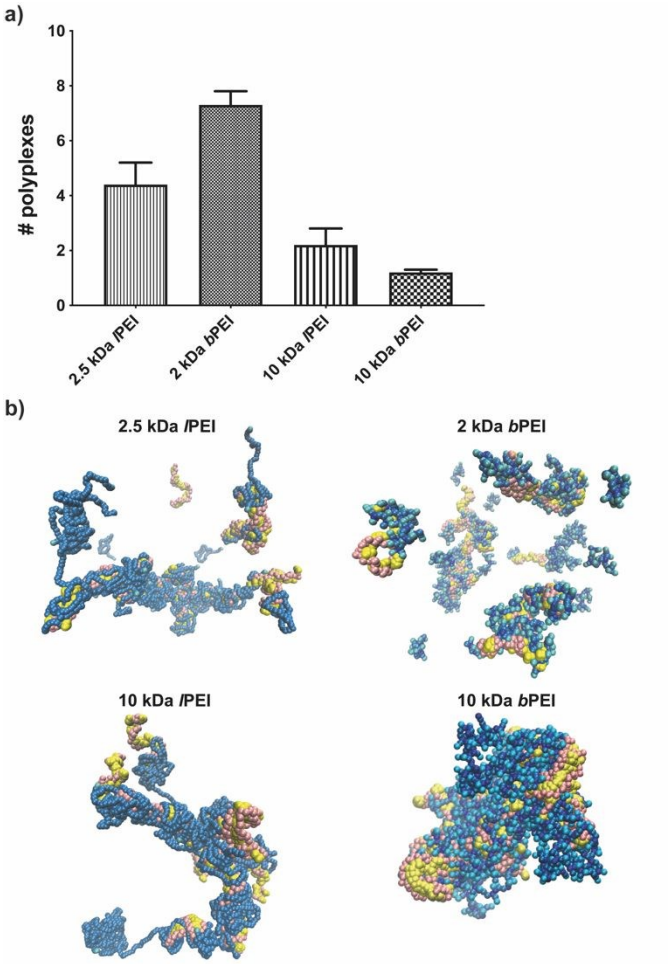


Figure 5. **a)** Number of siRNA-PEI polyplexes at N/P 10. LM_W PEI gave rise to the formation of some polyplexes, with siRNA molecules that were frequently singularly wrapped by PEI. This effect was particularly evident for 2 kDa *b*PEI. Conversely, in the case of 10 kDa PEIs, the formation of a limited number of considerably larger polyplexes that comprised almost all siRNA molecules in the system was apparent. Of note, in most of the simulation with 10 kDa *b*PEI, a single large cluster was observed. **b)** Visual depiction of siRNA-PEI polyplexes at N/P 10 as a function of the M_W of PEIs. LM_W PEIs lead to the formation of several small polyplexes that include 1 or 2 siRNA molecules. Conversely, 10 kDa M_W PEIs lead to the formation of a limited number of large clusters (from few to one) which comprised all siRNA molecules. Blue beads represent PEI monomers, yellow spheres are the nucleobases, while pink beads represent the ribose and the phosphate groups.

Conclusions

It has been recently shown that *in silico* molecular modelling is a powerful tool for research in biology, allowing to speed-up the design of therapeutic proteins,[67,68] enzymes,[69,70] and possibly solving the long-lasting protein folding problem with the very recent results obtained by AlphaFold in CASP14. By the same token, computational approaches are thought to aid in test benching a variety of transfectants and conditions, and speed up the development of ever more effective gene delivery

vectors. In this context, the goal of our work was to shed light on whether molecular simulations based on CG modelling might be useful to identify in the large variety of transfectants and conditions available, those which are the most suitable for complexing NAs and transfecting cells. This work is another step forward in the implementation of reliable tools which would help expedite the development and the testing of novel transfectants.

Herein, we have modelled polyplexes made of siRNA and relatively large M_W polycations of practical relevance. To best interpret our results from CG modelling and MD simulations and make them worthwhile for the scientific community, we identified a new index Q_N which takes into account the overall charge of the largest polyplex resulting from the simulations performed under different conditions. For any polycation tested, CG simulations showed that the electroneutrality of the largest polyplex was attained at the lowest N/P needed to achieve the full complexation of siRNAs. To the best of our knowledge, this is the first evidence that CG simulations can be used profitably as a predictive *in silico* tool to test experimentally-relevant transfectants and conditions, and to identify the optimum N/P for NAs complexation and gene silencing. In perspective, CG models may be used to design gene delivery vectors, saving wet-lab cost and time.

Of note, although an accurate depiction of the whole transfection process happening *in vivo* cannot be achieved computationally, CG and experimental results gave us reliable indications on the transfection behaviour of polyplexes *in vitro*. Specifically, the mixing of siRNA with relatively high M_w PEIs gives rise to the formation of only few, large polyplexes that are very effective in transfection. Instead, the addition of LM_w PEIs to NAs leads to numerous, smaller polyplexes, which are less transfection efficient. These findings are in agreement with the existing literature,[19] and are another step forward in shedding light on the way some specific features affect the transfection behaviour of gene delivery vectors.

Future work will focus on the validation of this integrated *in vitro-in silico* approach with other non-viral gene delivery vectors, such as other cationic polymers (dendrimers at different generations and *ad hoc*-synthetized polypeptides) and lipids. Nevertheless, with the increasing computational power and/or higher degree of coarse graining, NA larger than siRNA will become fully addressable.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was partly supported by Fondazione Cariplo (grant no. 2016-0481). The authors would like to thank Politecnico di Milano for financial support.

References

1 I. Lostalé-Seijo and J. Montenegro, *Nat. Rev. Chem.*, 2018, **2**, 258–277.
2 T. Wirth, N. Parker and S. Ylä-Herttuala, *Gene*, 2013, **525**, 162–9.
3 L. A. Jackson, E. J. Anderson, N. G. Rouphael, P. C. Roberts, M. Makhene, R. N. Coler, M. P. McCullough, J. D. Chappell, M. R. Denison, L. J. Stevens, A. J. Pruijssers, A. McDermott, B. Flach, N. A. Doria-Rose, K. S. Corbett, K. M. Morabito, S. O'Dell, S. D. Schmidt, P. A. Swanson, M. Padilla, J. R. Mascola, K. M. Neuzil, H. Bennett, W. Sun, E. Peters, M. Makowski, J. Albert, K. Cross, W. Buchanan, R. Pikaart-Tautges, J. E. Ledgerwood, B. S. Graham and J. H. Beigel, *N. Engl. J. Med.*, , DOI:10.1056/nejmoa2022483.
4 E. E. Walsh, R. W. Frenck, A. R. Falsey, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, K. Neuzil, M. J. Mulligan, R. Bailey, K. A. Swanson, P. Li, K. Koury, W. Kalina, D. Cooper, C. Fontes-Garfias, P.-Y. Shi, Ö. Türeci, K. R. Tompkins, K. E. Lyke, V. Raabe, P. R. Dormitzer, K. U. Jansen, U. Şahin and W. C. Gruber, *N. Engl. J. Med.*, , DOI:10.1056/nejmoa2027906.
5 L. Baldi, D. L. Hacker, M. Adam and F. M. Wurm, *Biotechnol. Lett.*, 2007, 29, 677–684.
6 T. M. Reineke, R. T. Raines and V. M. Rotello, *Bioconjug. Chem.*, 2019, **30**, 261–262.
7 P. L. Felgner, T. R. Gadek, M. Holm, R. Roman, H. W. Chan, M. Wenz, J. P. Northrop, G. M. Ringold and M. Danielsen, *Proc. Natl. Acad. Sci. U. S. A.*, 1987, **84**, 7413–7417.
8 R. W. Malone, P. L. Felgner and I. M. Verma, *Proc. Natl. Acad. Sci. U. S. A.*, 1989, **86**, 6077–6081.
9 N. Kose, J. M. Fox, G. Sapparapu, R. Bombardi, R. N. Tennekoon, A. Dharshan De Silva, S. M. Elbashir, M. A. Theisen, E. Humphris-Narayanan, G. Ciaramella, S. Himansu, M. S. Diamond and J. E. Crowe, *Sci. Immunol.*, , DOI:10.1126/sciimmunol.aaw6647.
10 F. Ponti, M. Campolungo, C. Melchiori, N. Bono and G. Candiani, *Chem. Phys. Lipids*, 2021.
11 W. T. Godbey and A. G. Mikos, *J. Control. Release*, 2001, **72**, 115–125.
12 Y. Rui, D. R. Wilson and J. J. Green, *Trends Biotechnol.*, 2019, 37, 281–293.
13 D. Pezzoli and G. Candiani, *J. Nanoparticle Res.*, 2013, 15.
14 M. Breunig, U. Lungwitz, R. Liebl and A. Goepferich, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 14454–14459.
15 W. T. Godbey, K. K. Wu and A. G. Mikos, *J. Control. Release*, 1999, 60, 149–160.
16 C. Malloggi, D. Pezzoli, L. Magagnin, L. De Nardo, D. Mantovani, E. Tallarita and G. Candiani, *Polym. Chem.*, 2015, **6**, 6325–6339.
17 C. Sun, T. Tang and H. Uludağ, *J. Phys. Chem. B*, 2012, **116**, 2405–2413.
18 H. Lee, A. K. R. Lytton-Jean, Y. Chen, K. T. Love, A. I. Park, E. D. Karagiannis, A. Sehgal, W. Querbes, C. S. Zurenko, M. Jayaraman, C. G. Peng, K. Charisse, A. Borodovsky, M. Manoharan, J. S. Donahoe, J. Truelove, M. Nahrendorf, R. Langer and D. G. Anderson, *Nat. Nanotechnol.*, 2012, **7**, 389–393.
19 D. Pezzoli, E. Giupponi, D. Mantovani and G. Candiani, *Nat.*

Publ. Gr., 2017, 1–11.

20 A. B. Hill, M. Chen, C. K. Chen, B. A. Pfeifer and C. H. Jones, *Trends Biotechnol.*, 2016, 34, 91–105.

21 T. G. Park, J. H. Jeong and S. W. Kim, *Adv. Drug Deliv. Rev.*, 2006.

22 N. Bono, F. Ponti, D. Mantovani and G. Candiani, *Pharmaceutics*, 2020.

23 M. Kim, H. R. Kim, S. Y. Chae, R. G. Larson, H. Lee and J. C. Park, *J. Phys. Chem. B*, 2013, **117**, 6917–6926.

24 J. Ziebarth and Y. Wang, *J. Phys. Chem. B*, 2010, **114**, 6225–6232.

25 Y. Nademi, T. Tang and H. Uludağ, *Nanoscale*, 2020, **12**, 1032–1045.

26 D. Meneksedag-Erol, J. N. Kizhakkedathu, T. Tang and H. Uludağ, *ACS Appl. Mater. Interfaces*, 2018, **10**, 28399–28411.

27 C. Sun and T. Tang, *J. Adhes. Sci. Technol.*, 2014, **28**, 399–416.

28 D. Meneksedag-Erol, T. Tang and H. Uludağ, *Biomaterials*, 2018, **156**, 107–120.

29 D. Meneksedag-Erol, T. Tang and H. Uludag, *J. Phys. Chem. B*, 2015, **119**, 5475–5486.

30 F. Merzel, F. Fontaine-Vive, M. R. Johnson and G. J. Kearley, *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, , DOI:10.1103/PhysRevE.76.031917.

31 J. Sponer, G. Bussi, M. Krepl, P. Banas, S. Bottaro, R. A. Cunha, A. Gil-Ley, G. Pinamonti, S. Poblete, P. Jurečka, N. G. Walter and M. Otyepka, *Chem. Rev.*, 2018, 118, 4177–4338.

32 Y. Z. Shi, F. H. Wang, Y. Y. Wu and Z. J. Tan, *J. Chem. Phys.*, , DOI:10.1063/1.4894752.

33 T. A. Knotts, N. Rathore, D. C. Schwartz and J. J. De Pablo, *J. Chem. Phys.*, , DOI:10.1063/1.2431804.

34 I. P. Kikot, A. V. Savin, E. A. Zubova, M. A. Mazo, E. B. Gusarova, L. I. Manevitch and A. V. Onufriev, *Biophysics (Oxf).*, 2011, **56**, 387–392.

35 K. W. Wang, K. Barker, S. Benner, T. Betancourt and C. K. Hall, *Mol. Simul.*, 2018, **44**, 1004–1015.

36 S. Mahajan and T. Tang, *J. Comput. Chem.*, 2019, **40**, 607–618.

37 J. D. Ziebarth and Y. Wang, *Biomacromolecules*, 2010, **11**, 29–38.

38 D. Meneksedag-Erol, T. Tang and H. Uludağ, *J. Phys. Chem. B*, 2015, **119**, 5475–5486.

39 J. D. Ziebarth, D. R. Kennetz, N. J. Walker and Y. Wang, *J. Phys. Chem. B*, 2017, **121**, 1941–1952.

40 T. A. Beu, A. E. Ailenei and A. Farcaș, *Chem. Phys. Lett.*, 2019, **714**, 94–98.

41 T. A. Beu, A. Ailenei and R. Costinaș, *J. Comput. Chem.*, 2019, jcc.26110.

42 S. Mahajan and T. Tang, *J. Phys. Chem. B*, 2019, **123**, 9629–9640.

43 K. A. Whitehead, R. Langer and D. G. Anderson, *Nat. Rev. Drug Discov.*, 2009, 8, 129–138.

44 T. J. Macke and D. A. Case, *ACS Symp. Ser.*, 1998, **682**, 379–393.

45 J. J. Uusitalo, H. I. Ingólfsson, S. J. Marrink and I. Faustino, *Biophys. J.*, 2017, **113**, 246–256.

46 L. Melone, B. Rossi, N. Pastori, W. Panzeri, A. Mele and C. Punta, *Chempluschem*, 2015, **80**, 1408–1415.

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

48 D. Van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, *J. Comput. Chem.*, 2005, **26**, 1701–1718.

49 F. Rigoldi, A. Gautieri, A. Dalle Vedove, A. P. Lucarelli, S. Vesentini and E. Parisini, *Proteins*, 2016, **84**, 744–758.

50 F. Rigoldi, P. Metrangolo, A. Redaelli and A. Gautieri, *J. Biol. Chem.*, 2017, **292**, jbc.M116.770271.

51 A. Gautieri, M. Ionita, D. Silvestri, E. Votta, S. Vesentini, G. B. Fiore, N. Barbani, G. Ciardelli and A. Redaelli, *J. Comput. Theor. Nanosci.*, 2010, **7**, 1–7.

52 A. Gautieri, S. Vesentini, A. Redaelli and M. J. Buehler, *Int. J. Mater. Res. (formerly Zeitschrift fuer Met.*, 2009, **100**, 921–925.

53 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, , DOI:10.1063/1.2408420.

54 H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. Dinola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.

55 S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman and A. H. de Vries, *J. Phys. Chem. B*, 2007, **111**, 7812–7824.

56 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33–38.

57 W. T. Godbey, K. K. Wu and A. G. Mikos, *J. Biomed. Mater. Res.*, 1999, **45**, 268–275.

58 D. W. Pack, A. S. Hoffman, S. Pun and P. S. Stayton, *Nat. Rev. Drug Discov.*, 2005.

59 a. Lucotti, M. Tommasini, D. Pezzoli and G. Candiani, *RSC Adv.*, 2014, **4**, 49620–49627.

60 A. L. Bolcato-Bellemin, M. E. Bonnet, G. Creusat, P. Erbacher and J. P. Behr, *Proc. Natl. Acad. Sci. U. S. A.*, , DOI:10.1073/pnas.0707831104.

61 A. P. Perez, E. L. Romero and M. J. Morilla, *Int. J. Pharm.*, , DOI:10.1016/j.ijpharm.2009.06.035.

62 M. Breunig, C. Hozsa, U. Lungwitz, K. Watanabe, I. Umeda, H. Kato and A. Goepferich, *J. Control. Release*, , DOI:10.1016/j.jconrel.2008.05.016.

63 Y. Hattori, A. Nakamura, S. Arai, M. Nishigaki, H. Ohkura, K. Kawano, Y. Maitani and E. Yonemochi, *Results Pharma Sci.*, , DOI:10.1016/j.rinphs.2014.01.001.

64 D. Pezzoli and G. Candiani, *J. Nanoparticle Res.*, 2013, **15**, 1523.

65 D. Pezzoli, E. Giupponi, D. Mantovani and G. Candiani, *Nat. Publ. Gr.*, 2017, 1–11.

66 V. Forest and J. Pourchez, *Mater. Sci. Eng. C*, 2017.

67 P. Chatterjee, M. Ponnapati, C. Kramme, A. M. Plesa, G. M. Church and J. M. Jacobson, *Commun. Biol.*, 2020, **3**, 1–8.

68 L. Cao, I. Goreshnik, B. Coventry, J. B. Case, L. Miller, L. Kozodoy, R. E. Chen, L. Carter, A. C. Walls, Y. J. Park, E. M. Strauch, L. Stewart, M. S. Diamond, D. Veesler and D. Baker, *Science*, , DOI:10.1126/science.abd9909.

69 F. Rigoldi, S. Donini, A. Redaelli, E. Parisini and A. Gautieri, *APL Bioeng.*, 2018, **2**, 011501.

70 Y. Wang, J. Chen and Z. Kang, *Biochemistry*, 2019.

Journal Name ARTICLE

View Article Online
DOI: 10.1039/D0NR09052B