JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

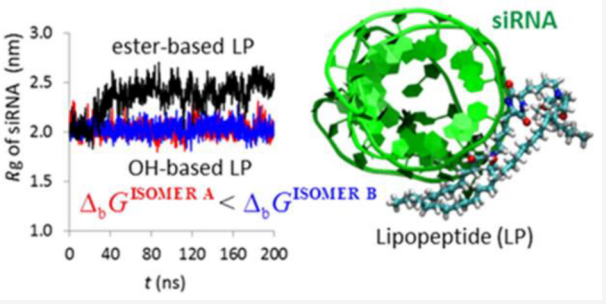
Binding Dynamics of siRNA with Selected Lipopeptides: A Computer-Aided Study of the Effect of Lipopeptides' Functional Groups and Stereoisomerism

Evangelia Pantatosaki* and George K. Papadopoulos*

Cite This: J. Chem. Theory Comput. 2020, 16, 3842−3855 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The engineering issues pertaining to nanoparticle systems toward targeted gene therapies have not been fully probed. Recent experiments have identified specific structural characteristics of a novel class of lipopeptides (LP) that may lead to potent nanocarriers intended as RNAi therapeutics, albeit the molecular mechanism that underlies their performance remains unexplored. We conducted molecular dynamics simulations in atomistic detail coupled with free energy computations to study the dynamics and thermodynamics of an acrylate- and an epoxide-derived LP, members of the aforesaid class, upon their binding to siRNA in aqueous solution aiming at examining structure-potency relations. We found that the entropic part of the free energy of binding predominates; moreover, the first LP class tends to disrupt the Watson−Crick base pairing of siRNA, whereas the latter leaves the double helix intact. Moreover, the identified tug-of-war effect between LP−water and LP−siRNA hydrogen bonding in the supramolecular complex can underpin synthesis routes toward tuning the association dynamics. Our simulations on two diastereomers of the epoxide-derived LP showed significant structural and energetics differences upon binding, as a result of steric effects imposed by the different absolute configurations at their chiral centers. These findings may serve as crucial design parameters toward modulating the interplay between complex stability and ease of releasing the nucleic acid drug into the cell.



INTRODUCTION

Over the last years, there has been rapidly increasing interest toward the development of RNA interference (RNAi) therapeutics by means of delivering to the cytoplasm of the target cells synthetic fragments of genetic material, known as short interfering RNAs (siRNA), to be incorporated into the RNAi natural pathway in order to prevent the production of disease-related proteins.[1,2] RNAi is an endogenous cellular process for regulating gene expression and is considered one of the most important recent discoveries in molecular oncology, recognized by a Nobel Prize in 2006.[3] However, the delivery of naked siRNA molecules has proved quite ineffective mainly because of their susceptibility to degradation by extracellular ribonucleases[4] and their likelihood of activating the innate immune system.[5]

The hurdle in harnessing the RNAi process is mainly attributed to the difficulty in discovering selective and efficient ways to guide the siRNA molecules inside the pathogenic cells. The development of various chemical modification routes has increased the resistance of siRNAs against endogenous nucleases and decreased the likelihood of inducing innate immune responses.[6] However, the permeation of siRNA molecules across the cellular lipid membrane remains an immense point at issue as a consequence of their large molecular weight (approximately 14 kDa) and their anionic nature (net negative charge).

To surmount these challenges, systemic delivery of siRNA up to the final site of action in the target cells has been designed over the last years, starting either via viral vectors[7] whose use is subject to obvious risks or by engineering nonviral delivery platforms. The latter class was based on cationic and/or ionizable lipids that self-assemble in micelles or liposomes, which entrap the negatively charged nucleic acids to enable efficient *in vivo* delivery. Other nonviral vectors exhibiting promising *in vivo* potency followed, such as dynamic polyconjugate polymer systems, trivalent targeting ligand−siRNA conjugates, and dendrimers. In refs 8−10 the reader can find extensive reviews on the use and design of such delivery systems. Notably, the type of nonviral vectors has been shown to affect the siRNA biodistribution in different parts of the body.[11,12] Recently, the FDA approved the first ever RNAi drug encapsulated in a lipid nanoparticle, which targets the so-

Received: December 18, 2019
Published: April 23, 2020

https://dx.doi.org/10.1021/acs.jctc.9b01261
J. Chem. Theory Comput. 2020, 16, 3842−3855

called hereditary transthyretin-related amyloidosis that can impair heart and nerve function.[13] This revolutionary advancement in the RNAi field yields hope for cure and paves the way for vectors that will in addition target several human organs beyond the liver wherein nanoparticles tend to accumulate upon systemic delivery.

The lipid nanoparticles (NP) are commonly constructed by the self-assembly of a mixture of distearoylphosphatidylcholine (DSPC), cholesterol, PEG-lipid molecules, and a key lipid variant, to form spherical entities encapsulating the siRNA molecules. As lipid variants, many compounds have been tested: optimized ionizable cationic lipids,[14] lipidoids comprising amino-alkyl-acrylate and -acrylamide materials,[15] or lipopeptide compounds formed by conjugating lipid tails to amino acids, peptides, and polypeptide head groups.[16]

However, the core−shell structure in molecular detail of these NPs remains unknown. Experiments have demonstrated that even a small chemical change to the molecular structure of the lipid component affects the NP silencing efficiency and cell viability both *in vitro* and *in vivo*. Although experiments have identified specific structural characteristics of the lipid components that may lead to potent NPs (e.g., type of head groups, head−tail linking groups, length and modification of the tails), the molecular mechanism behind their beneficial effects remains unexplored.[14−16] In addition, the structure, self-assembly, dynamics, and binding strength of the key lipid−siRNA complexes have to be examined in detail to link the lipid component's chemical structure with the NP performance.

In a manner complementary to conventional experiments, statistical mechanics-based computer "experiments" can shed light on the aforesaid characteristics by providing insights into atomistic detail of these oligonucleotide carrying systems. To this end, a number of modeling studies have been reported so far addressing siRNA carriers that include poly(amidoamine)[17−19] and triazine[20] dendrimers, as well as UV-degradable dendrons,[21] poly(L-lysine),[18] polyethylenimine[22,23] and PMAL[24] polymers, and arginine-rich peptides.[25]

In this work, we performed full-atom molecular dynamics simulations to model the supramolecular association of NP's lipid molecules with siRNA, aiming at exploring the thermodynamics and molecular dynamics of a lately appeared class of lipopeptides (LP). In particular, we investigated the mechanism of binding of siRNA with two LP members of this class, an acrylate-derived and an epoxide-derived, which have exhibited promising *in vitro* and *vivo* results in gene silencing experiments[16] in HeLa cells, rodents, and nonhuman primates, characterized by different functional groups in their structure, and bearing multiple chirality centers. Our objective is to understand the structure−potency relation and the effect of stereoisomerism on these LP−siRNA complexes, with respect to their observed[16] potency as components of nanoparticle systems.

Stereoisomerism is of major importance in biochemical and pharmaceutical processes;[26,27] nucleic acids, and hence siRNAs, exhibit inherent chirality (right-handed double helices) and comprise chirality centers in the nucleotides.[28] Pure enantiomers of pharmaceutical agents that target the DNA double helix have been shown to exhibit different potency in sequence-specific binding[29] and chemotherapeutic efficacy.[30] Since the synthesis reactions of the LPs of this study exhibit no stereoselectivity, a mixture of stereoisomers is present in the NPs' formulations. To address the aforementioned experimental issues, we proceeded with the modeling of two different stereochemical forms of the epoxide-derived lipopeptide. To the best of our knowledge, this is the first time that stereoisomerism is addressed, either computationally or experimentally, in a siRNA-carrier system.

The work is organized as follows: Full-atom computer reconstructions of the novel LPs of this study, their stereoisomers, and the double stranded siRNA molecule are presented in the first part. Subsequently, molecular modeling details are given for all molecular dynamics of single and multi LP−siRNA systems at various ratios; binding configurational free energy and entropy calculations follow. The impacts of the lipopeptides' chemical composition and stereoisomerism on the binding mechanism and strength of the LP−siRNA complexes are discussed in the last section of this article along with the stability and conformational changes of the biomolecules upon complexation, their critical contacts and binding sites, and the intermolecular LP−siRNA and LP−H_2O hydrogen bond pairing.

MOLECULAR SIMULATIONS

Materials Modeling. The lipopeptides cKK-E12 (two diastereomers) and cKK-O12, both members of a lately synthesized lipopeptide class, were computer reconstructed by means of an in-house code, using as input structural data from ref 16. The generated Cartesian atomic coordinates and all intraconnectivity data (bonds, angles, dihedrals) were transcribed to a topology file that in turn was used as input in the simulations. The initial extended reconstructed structures of the LPs can be seen in Figure 1.

Both LPs comprise a 2,5-diketopiperazine core, namely, a cyclic dipeptide (dilysine-derived) head core (HD) that consists of a heterocyclic six-membered ring containing two polar amide

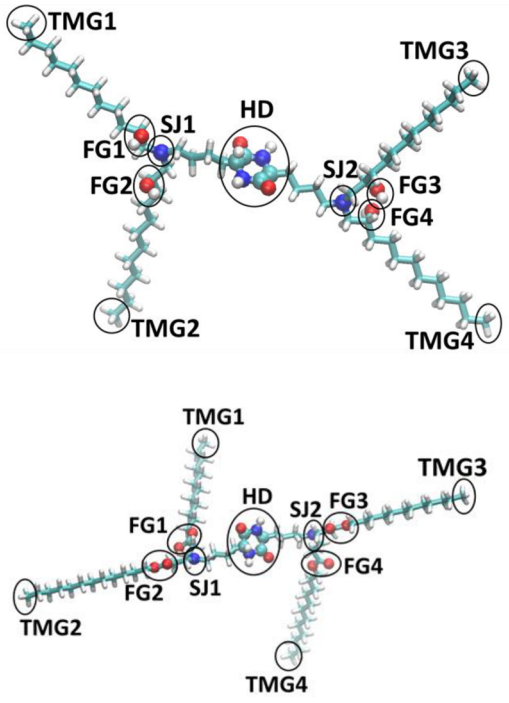


Figure 1. Extended conformations of the reconstructed lipopeptides; cKK-E12 (top) and cKK-O12 (bottom) to depict the head core (HD), shoulder joints (SJ), and lipid tails with the functional groups (FG) and terminal methyl groups (TMG). Color code: nitrogen (blue), oxygen (red), carbon (cyan), hydrogen (white).

https://dx.doi.org/10.1021/acs.jctc.9b01261
J. Chem. Theory Comput. 2020, 16, 3842−3855

groups (Figure 1). Two hydrocarbon branches adhere to the HD ending to two tertiary amine groups, termed here as shoulder joints (SJ); two lipid tails branch off from each SJ. The lipid tails are chemically modified via functional groups (FG) at positions located near the SJs. The methyl groups located at the free end of the tails are designated here as terminal methyl groups (TMG).

The LPs bear chemically different lipid tails, that is to say, alcohol-based (1,2-epoxydodecane-derived) for cKK-E12, and ester-based (dodecyl acrylate-derived), for cKK-O12. The molecular formulas are $C_{60}H_{120}N_4O_6$ for cKK-E12 and $C_{72}H_{136}N_4O_{10}$ for cKK-O12 (details in Figure S1). We also reconstructed the protonated forms of the molecules since the complexation with siRNA is carried out *in vitro* under acidic conditions, where half of the LP tertiary amines are protonated; in both LPs it is the SJ1 nitrogen atom that was protonated (Figure 1 and Figure S1).

Because both LP classes possess multiple chiral centers, the different spatial arrangement of the substituents on a chiral atom, i.e., absolute configurations, gives rise to different stereoisomers. The cKK-O12 compound has two asymmetric carbons located on HD where the hydrocarbon branches adhere (Figure S1). The cKK-E12 molecule has the same two asymmetric carbons as cKK-O12, and, in addition, four asymmetric carbons located at the tails where the hydroxyl FGs adhere (Figure S1).

cKK-O12 was modeled in its absolute configuration *S*, *R* (at carbons on HD near the SJ1 and SJ2, respectively). Two diastereomers of cKK-E12 were modeled, coded here as (i) cKK-E12-A with absolute configurations *S*, *R* (at carbons on HD near the SJ1 and SJ2) and *R* (four carbons at the tails); and (ii) cKK-E12-B with absolute configuration *S* (at all six asymmetric carbons). It is noted that cKK-E12-A and cKK-O12 have the same absolute configurations at the HD chirality centers. The *R* or *S* descriptors of the absolute configuration follow the Cahn−Ingold−Prelog convention.[31]

The short interfering RNAs are duplexes of 21−23 nucleotides, approximately 8 nm in length and 2 nm wide.[32] The initial structure of the siRNA was taken from X-ray diffraction data deposited in the protein data bank (PDB code: 2F8S);[33] it is a 22-mer siRNA with two nucleotide 3′ overhangs at the two ends of the strands. The sequence of bases along the two strands can be found in Table SI of the Supporting Information.

Force fields and molecular dynamics. We performed molecular dynamics (MD) simulations to record the time evolution of the LP configurations under isobaric−isothermal conditions (*N*, *P*, *T*), where *N* is the total atoms involved in the systems under study: (i) isolated LP; (ii) protonated LP in aqueous solution; (iii) LPs dissolved in water and ethanol; (iv) protonated LPs with one siRNA molecule in aqueous solution at the molecular ratios LP:siRNA of 1:1, 4:1, and :1; and (v) one naked siRNA in aqueous solution simulated for the sake of comparison. In the above systems wherein siRNA is involved, randomly selected water molecules were replaced by Na cations to neutralize the negative charge on its phosphate backbone. Also, the appropriate number of Cl anions was added similarly to neutralize the protonated amine sites of the lipopeptides. The final solvated system contained approximately 224 000 atoms for the largest system, 8:1. MD trajectories were produced up to the max time length of 350 ns.

The basic computations of the systems were carried out by setting the algorithm thermostat at 310 K. In addition, simulations for T = 290 K and T = 300 K were performed for the protonated LPs in aqueous solution and the LP:siRNA ratio of 1:1 to compute the Gibbs energy of binding for the LP−siRNA complex as a function of temperature, thereby the binding entropy of the complex was extracted (section D).

Then, numerical integration of the equations of motion[34] for all the degrees of freedom involved in the systems was performed by employing proper force fields for the lipopeptides and the siRNA molecules. In particular, in the absence of an *ad hoc* force field for lipopeptides we opted for AMBER99sb-ildn,[35] which has been successfully employed in peptide−nucleic acid simulations;[36,37] the employed partial charges on the LP atoms were taken from OPLS-AA[38,39] (the way the LP atoms have been assigned to the specific force fields' atom types, which appears in Tables SII and SIII and Figure S1b). For the siRNA molecule the AMBER ff94[40] force field was adopted; this force field is one of the most widely used atomistic force fields for modeling nucleic acids and is incorporated in most popular software packages[34,41] for biomolecular simulations.

It should be noted that early modifications of the AMBER ff94 force field, by means of the ff98[42] and ff99[43] versions have brought about slight changes to the χ torsion angle and sugar pucker parameters; however, these variants perform rather similarly to the original ff94.[44−46] Subsequent AMBER reparametrizations as found in parmbsc0[47] have improved the performance of B-DNA simulations through a reparametrization of the α/γ dihedral angles, albeit they do not substantially affect simulations of the RNA canonical A-form duplexes.[44,48]

A prominent AMBER modification focusing specifically on improving RNA modeling is the ff99bsc0χ_{OL3} variant[48,49] that accurately captures and maintains stable RNA conformations for microsecond-long simulations. Two other recent AMBER modifications for RNA include (1) a reparametrization[50] combined with a specific four-site water model[51] and (2) a variant using the addition of a correction term to the AMBER potential function to selectively tune the hydrogen-bonding interactions.[52] It is also worth mentioning the recently appearing OPLS-AA/M[53] that reports a careful refinement of the torsional parameters by means of high-level DFT calculations and NMR experimental data, exhibiting significant improvements in the description of dinucleotides, tetranucleotides, and longer sequences with noncanonical RNA motifs. Nevertheless, none of the above-mentioned force fields have been incorporated so far in the official releases of the GROMACS software package[34,54] that we used for the integration of the multiple degrees of freedom of the systems under study. For a thorough review on RNA force fields the reader is referred elsewhere.[44−46,55]

Special care should be taken in the choice of the water model, since RNA simulations using different AMBER variants have proved to be water-model sensitive.[45,46,56,57] In our study the AMBER ff94 was combined with the three-site TIP3P water model,[58] which enables high computational efficiency and has been successfully used with AMBER force fields in modeling RNA hydration.[57] In the next paragraphs, simulation issues on the systems' setup, equilibration, and production runs are given.

For the siRNA systems we used a rhombic dodecahedron simulation box to allow solvent to surround the nucleic acid, excluding at the same time large regions containing water molecules to avoid a dramatic increase of computational cost. The box dimensions were set to ensure at least 3 nm thickness for the solvent surrounding the LP−siRNA system in all directions. This value is well above the minimum value of 0.8−

https://dx.doi.org/10.1021/acs.jctc.9b01261
J. Chem. Theory Comput. 2020, 16, 3842−3855

1.2 nm recommended in current protocols for nucleic acids simulations.[59] Furthermore, it ensures both the presence of the hydration shell, identifiable in the double stranded nucleic acids, and the necessary amount of surrounding bulk water.[60] The total number of water molecules in the box was over 73 600 for the largest system of this study (LP:siRNA at the ratio of 8:1).

Energy minimization prior to MD production runs was performed for all systems by means of the steepest-descent numerical method to eliminate unphysical atom overlap. During minimization, the LPs and siRNA were fixed at their initial conformations using harmonic restraints with a force constant of 10^6 J mol^{-1} nm^{-2}. Subsequently, equilibration MD runs were started under the isothermal−isochoric ensemble (N, V, T) for a time length of 2 ns to stabilize the temperature of the solvent at the prescribed temperature, followed by 2 ns in the (N, P, T) ensemble, setting P = 1 bar to stabilize pressure and density of the surrounding medium. Then, the imposed constrictions on the macromolecules were removed and MD production runs were followed up to 200 ns. The Parrinello−Rahman algorithm was utilized to allow for fluctuating the pressure tensor (diagonal components) isotropically under the constraint of a constant trace (the pressure),[61] combined with velocity-rescaling[62] for the thermostating. Bond lengths were constrained as usually in most biomolecular systems.[63] For the computation of electrostatic interactions the particle-mesh Ewald method[64] was employed. The produced trajectories for postprocessing were being recorded every 10 ps. Visualization of all trajectories and image rendering were performed ultimately by means of VMD.[65]

Thermodynamics of Binding. The excess molar Gibbs energy difference of solvation in units of energy per mol of solute,

$$\Delta_S G_e \equiv G_e(T, P, \{x\}) - G(\varepsilon = 1) \tag{1}$$

namely, the reversible work, $\frac{1}{N_A}\Delta_S G_e$, performed under constant pressure and temperature, where N_A is the Avogadro constant when an ion is transferred from vacuum (dielectric constant ε = 1) to a solvent of dielectric constant ε and density ρ, consisting of n species with $\{x\}$= $\{x_1, ..., x_{n-1}\}$ molar fractions, is given by the following relation.[66]

$$\Delta_S G_e = \int_0^q dq' \int \mathbf{dr}\, r^{-1} \sum_l h_l(r; q')\rho q_l \tag{2}$$

For the derivation of eq 2, a "growing" procedure over a continuum spectrum of charges up to the value q was followed,[67] with $h_l(r, q')$ being the equilibrium distribution function around a solute with charge q'; the subscript, l, stands for the charge sites on solvent molecules. It is noteworthy that a **k**-space transformation of the above spatial integral, assuming an effective Born radius independent of the solute charge (not valid for small solutes bearing large charges),[68] makes eq 1 consistent with the linear dielectric response realization of Born excess molar free energy of hydration for spherical ions of radius σ_e, based on classical electrostatics, $\Delta_S G_e = \frac{q^2}{2\sigma_e}\left(\frac{1}{\varepsilon} - 1\right)$.[69]

The electrostatic interaction energy, $\mathcal{V}_e \equiv \mathcal{V}_e(q)$, between a solute of charge q and the solvent can be evaluated through the same spatial integral of eq 1 to give an average of the form

$$\langle \mathcal{V}_e \rangle = \int \mathbf{dr}\, q r^{-1} \sum_l h_l(r; q)\rho q_l \tag{3}$$

Further, a relation between the Gibbs energy of solvation and the average solute−solvent electrostatic energy, initially appearing as an empirical result[70] that later was demonstrated by a rigorous statistical mechanics formulation,[68] gives

$$\Delta_S G_e = \frac{1}{2}\langle \mathcal{V}_e \rangle \tag{4}$$

Equation 4 offers a convenient way of computing in atomistic simulations the electrostatic part of the solvation free energy difference; for the remaining part, that is to say, the nonpolar contribution, namely, the van der Waals interactions potential, $\mathcal{V}_d$, we adopted a linear approximation[71] similar to eq 4 for $\Delta_S G_d$ from $\langle \mathcal{V}_d \rangle$ (see section D), properly adapted here for the study of the lipopeptide−siRNA complexation system. Therefore, using the notation $\Delta_S G_t$ for the total molar free energy difference of solvation, we obtain, analogously to eq 1, the following equations for the two systems: LP in water (system 1), $\Delta_S G_t^{(1)} = G_t^{(1)}(T, P, \{x\}) - G(\varepsilon = 1)$, and LP in water plus siRNA (system 2), $\Delta_S G_t^{(2)} = G_t^{(2)}(T, P, \{x\}) - G(\varepsilon = 1)$.

Both quantities can be evaluated by eqs 5 and 6 in the course of two (N, P, T) MD simulations, in the presence of the respective counterions, namely, protonated LP plus chlorine ion and siRNA plus sodium ions. Then, the Gibbs energy of binding, $\Delta_b G$, was computed for the complexation process

$$[\mathrm{LP}]_{aq} + [\mathrm{siRNA}]_{aq} \rightarrow [\mathrm{LP} \cdots \mathrm{siRNA}]_{aq}$$

Partial charges and counterions have been dropped for simplicity in the above reaction scheme. Because energy is a state function, $\Delta_b G$ is

$$\Delta_b G \equiv G_t^{(2)}(T, P, \{x\}) - G_t^{(1)}(T, P, \{x\}) \tag{5}$$

or

$$\Delta_b G = \frac{1}{2}(\langle \mathcal{V}_e^{(2)} \rangle - \langle \mathcal{V}_e^{(1)} \rangle) + \alpha(\langle \mathcal{V}_d^{(2)} \rangle - \langle \mathcal{V}_d^{(1)} \rangle) \tag{6}$$

where α is an empirical parameter to be determined via calibration from experimental data, as detailed in the section D. Calculation of the molar entropy of binding for the aforementioned closed systems can be done by recalling macroscopic thermodynamics through the following relation:

$$\left(\frac{\partial \Delta_b G}{\partial T}\right)_{P,\{x\}} = -\Delta_b S \tag{7}$$

by repeating the described simulations at different temperatures under the same pressure.

Taking the derivative of eq 7 with respect to temperature under constant pressure and mass, by virtue of the fundamental equation for the Gibbs energy, $\Delta_b G = \Delta_b H - T\Delta_b S$, we get the temperature dependence of $\Delta_b S$ via eq 8 as

$$\left(\frac{\partial \Delta_b S}{\partial T}\right)_{P,\{x\}} = -\left(\frac{\partial^2 \Delta_b G}{\partial T^2}\right)_{P,\{x\}} = \frac{C_P^{(2)} - C_P^{(1)}}{T} \tag{8}$$

with C_P being the molar heat capacity under constant pressure.

RESULTS AND DISCUSSION

A. Lipopeptides Modeling. Molecular dynamics simulations of a single isolated anhydrous LP in the canonical Ensemble (N, V, T), where T = 310 K, and V is the primary simulation box volume, were initially executed for 1 ns using as input configurations the extended structures illustrated in Figure

https://dx.doi.org/10.1021/acs.jctc.9b01261
J. Chem. Theory Comput. 2020, 16, 3842−3855

1. The computer runs were carried out both with and without bond constraints leading to insignificant differences. Subsequently, the equilibrated LP structures were modeled in their protonated forms surrounded by water molecules plus a chlorine anion to neutralize the system. The MD trajectories (310 K, 1 bar) were traced for a time length of 200 ns. We found that the hydrated LP structures obtain preferably folded conformations; the equilibrated configurations are dictated by the competition between the propensity of (a) the head, (b) the protonated tertiary amine, (c) the functional groups, for the solvent, and (d) the tendency of lipid tails to minimize solvent contact. This finding resembles the behavior of amphiphilic surfactant molecules in aqueous solutions found earlier by other researchers.[72]

Radii of gyration, R_g, calculated over the entire LP trajectories are shown in Figure 2; the considerably higher R_g of the initial

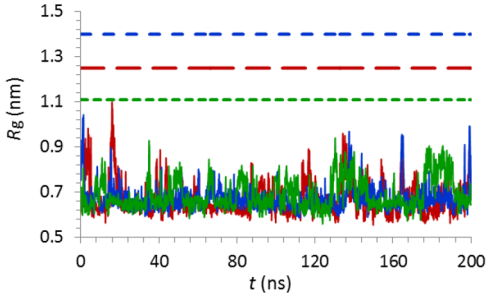


Figure 2. Time evolution of the calculated radius of gyration of the protonated cKK-O12 (blue), and diastereomers cKK-E12-A (red) and cKK-E12-B (green) in aqueous solution at 310 K and 1 bar over the entire MD trajectories. Lines indicate calculated R_g values of extended conformations for comparison.

extended conformations are also depicted for the sake of comparison. The calculated R_g for all hydrated LP structures exhibit similar average values and fluctuations, with the observed compaction to be greater for cKK-O12 with respect to its extended conformation, i.e., 52% for cKK-O12, 46% for cKK-E12-A, and 39% for cKK-E12-B. This result may be attributed to the length of the LP hydrocarbon tails and also the type of functional groups existing at the tails, as discussed in the next paragraph.

Postprocessing of the equilibrated MD trajectories showed that the HBs formed by the OH groups with water are multiple and stronger compared to the HBs formed by the cKK-O12 ester groups with water; the computed number of HBs, as well the HB length and angle distributions appear in Figure S2. The tails of the cKK-E12 diastereomers appear to be less hydrophobic due to the presence of the polar hydroxyl FGs. It should be noted here that for classifying a bond as HB the IUPAC's criteria[73] were followed (details in the Supporting Information). In addition, it is observed that both diastereomers form a comparable number of HBs of the same strength as water (Figure S3). Moreover, the presence of ester groups at the cKK-O12 longer hydrophobic tails seems to offer higher flexibility and facilitates folding (in analogy to incorporating ester linkages in a hydrocarbon polymer's main chain, which induces torsional mobility and structural flexibility[74,75]).

In view of the current first attempt to model these new LP structures, we simulated the mixture molecular dynamics in the (N, P, T) ensemble of the LPs dissolved in water, and ethanol, at T = 310 K and P = 1 bar for 100 ns, starting from a random uniform initial configuration including 32 LPs and 1280 solvent molecules (solute:solvent 1:40) in a cubic box (details in Supporting Information). In the first case spontaneous formation of a two-phase system was observed, while in the latter a homogeneous one-phase region (Figure S4) indicates the LPs' miscibility to ethanol. Also, in the ethanol solution both LP types are less compacted due to favorable interactions with the amphiphilic ethanol molecules; this is reflected in the increase of the computed R_g values of the LPs dissolved in ethanol vs water (Figure S5). It is worth mentioning also that the computed length and angle distributions (Figure S6) of the HBs that the LPs form with the solvent molecules show shorter HBs with smaller angles for the epoxide- vs the acrylate-derived LPs in both solvents, indicating higher bond strength for the former LP class, presumably due to the presence of polar OH FGs in its structure.

B. Supramolecular Complexation Issues. In this section we investigate the lipopeptide−siRNA association dynamics, the critical contacts and probable binding sites, as well as the conformational changes of the molecules upon binding, and the combatting of the intermolecular LP−siRNA and LP−H_2O hydrogen bond pairings. The calculations here are concerned with cKK-O12 and cKK-E12-A, both having the same absolute configurations at the HD chirality centers. The effect of the lipopeptides' stereoisomerism on the complexation mechanism is examined in section C.

Binding Dynamics. We monitored the time evolution of the minimum distance of every atom on siRNA with all atoms on the protonated LP biomolecules during an MD run for all LP:siRNA ratios using the following convention: distances up to a max value of 0.4 nm denote complexation. We found that for systems of higher ratios, a few LPs may stay unbound even up to 30 and 10 ns for cKK-E12-A and cKK-O12, respectively (Figure S7); the above behavior indicates a slower complexation dynamics for the former LP. Eventually, complexation with the siRNA was accomplished for all LPs in all systems.

We also computed the radii of gyration, R_g, of the lipopeptides (Figure S8, Table SIV) and the time evolution of the number of close contacts (NoCC) between LP molecules and siRNA (Figure S9), in the single and multi-LP systems along the entire MD trajectory. The NoCC is defined here as the number of the LP atoms whose distance from every atom on siRNA is less than 0.3 nm. As seen in Figure 2 and Figure S8, an increase of the LPs' radii of gyration after they have bound to siRNA is observed compared with results for their free hydrated states; both LPs were found to form multiple close contacts with siRNA (Figure S9). At the LP:siRNA ratio of 1:1, cKK-O12 exhibits higher NoCC values than cKK-E12-A (Figure S9), indicating a higher number of interaction sites exposed to siRNA. At the LP:siRNA ratio of 8:1 the NoCC for cKK-E12-A takes the highest value; this is also shown in Figure 3 wherein the time evolution of the NoCC for the cluster of eight LPs is depicted. In addition, computation of the R_g of the eight LPs as a single cluster results in a larger value for the cKK-E12-A cluster (vs cKK-O12), indicating that the participating molecules obtain a sparser layout along the biomacromolecule's surface.

This finding was further investigated by computing the number of water molecules that reside in the siRNA hydration shell (up to 0.35 nm) for all LP:siRNA ratios (Figures 3 and Figure S10). The results show that the number of water molecules in the shell decreases after complexation, especially at higher LP concentrations. Because there is no structural compaction of the siRNA as indicated from its R_g value in

https://dx.doi.org/10.1021/acs.jctc.9b01261
J. Chem. Theory Comput. 2020, 16, 3842−3855

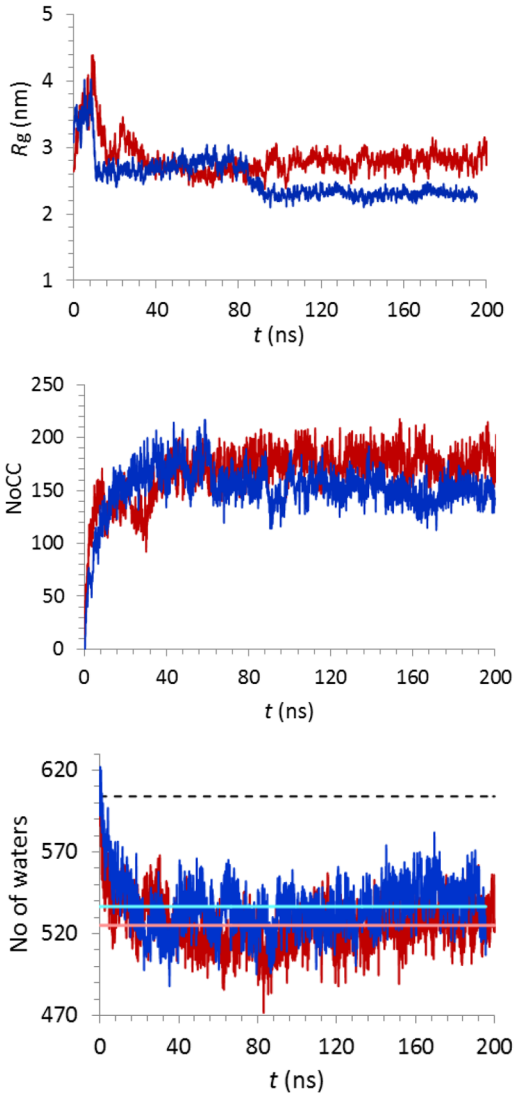


Figure 3. Time evolution of the following quantities for cKK-E12-A (red) and cKK-O12 (blue) at the 8:1 LP:siRNA ratio: radius of gyration of the cluster of eight LPs (top), number of close contacts between siRNA and the cluster of eight LPs (middle), and number of water molecules in the siRNA hydration shell (bottom). Continuous lines denote time averages; the dashed line denotes the number of water molecules in the hydration shell for the naked siRNA.

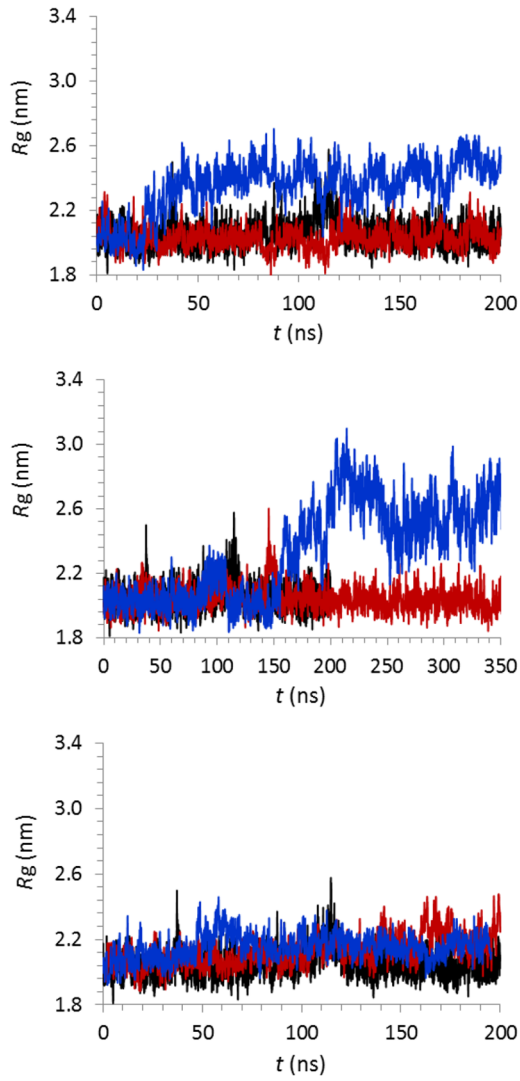


Figure 4. Time evolution of the siRNA radius of gyration upon binding with cKK-E12-A (red) and cKK-O12 (blue), at 1:1 (top), 4:1 (middle), and 8:1 (bottom) LP:siRNA ratios. The R_g of the naked siRNA in aqueous solution (black) also appears.

Figure 4, the observed loss can be explained on the basis of water molecules expelled from the shell by the LPs at the proximity of the double helix. It is also observed that the number of water molecules ejected from the shell is higher in the case of cKK-E12-A than for cKK-O12; this is particularly true at high LP concentrations, as shown in Figure 3, indicating that the cKK-E12-A molecules occupy a larger area along the surface of the oligonucleotide.

Critical Binding Sites. In all LP:siRNA systems, LPs were placed initially at random positions in the simulation box at maximum distances about 2 nm from the siRNA. They were found to bind preferentially to the minor groove of the siRNA, as revealed by the visualization of the entire MD trajectory (Figure S11). To elucidate the binding motif of the LPs to siRNA, identification of the critical binding sites on the LP structure was carried out by means of computing the minimum distance between important chemical moieties of the LPs, such as SJs, FGs, and TMGs with the siRNA, at LP:siRNA of 1:1, as shown in Figure S12. The minimum distance shown is the minimum value among the distances computed for every atom of each of the above LP moieties with all atoms of siRNA. Both LPs form permanent close contacts with the siRNA through their protonated tertiary amine group SJ1 (Figure S12). Such a strong binding of the protonated tertiary amine group to the nucleic acid resembles the hydrogen bond pairing of the lysine side chains of proteins to DNA backbone atoms in protein–DNA complexes, identified by X-ray crystallography.[76]

In the same figure (Figure S12), the functional groups FG1 and FG2, which are close to the protonated SJ1, bind closely to siRNA, preserving smaller distances compared to those for FG3 and FG4, which are close to the neutral tertiary amine group SJ2. The latter FGs preserve longer distances from the siRNA, indicating occasionally unbound FGs. The hydrophobic TMGs of both molecules form labile contacts being frequently further away from the siRNA; the shown fluctuations are larger for the TMGs on cKK-E12-A, reflecting higher mobility of its tails. Furthermore, the hydrogen atoms of the HD amide groups were

https://dx.doi.org/10.1021/acs.jctc.9b01261
J. Chem. Theory Comput. 2020, 16, 3842−3855

found to form close contacts with the O2′ ribose sugar atom of the siRNA (Figure S13).

Conformational Stability of the siRNA. The time evolution of the siRNA radius of gyration after complexation is depicted in Figure 4. It is observed that at LP:siRNA of 1:1 and 4:1 the binding with cKK-O12 raises the R_g of the siRNA; analogous behavior has been reported in the case of PAMAM dendrimer carriers.[17] This increase in R_g becomes less pronounced at the highest complex ratio, cKK-O12:siRNA of 8:1. By contrast, we found that within the computation times of this study the cKK-E12-A complex forms without noticeable alteration of the siRNA gyration radius.

To further explore the conformational stability of the siRNA double helix, we computed the distance between the stacked bases in each siRNA strand and also the distance between opposing bases in a base pair (bp), for the ratio 1:1. These calculations include only bases belonging to the nucleotides whereat the LP binds (details in the Supporting Information, Figure S15). The way the bases are numbered in each siRNA strand and base pairs in the double helix are explained in Table SI.

Our computations show that for cKK-O12, the distance between some of the stacked bases increases, as shown for one strand (strand I in Table SI) of the siRNA in Figure 5 (similar

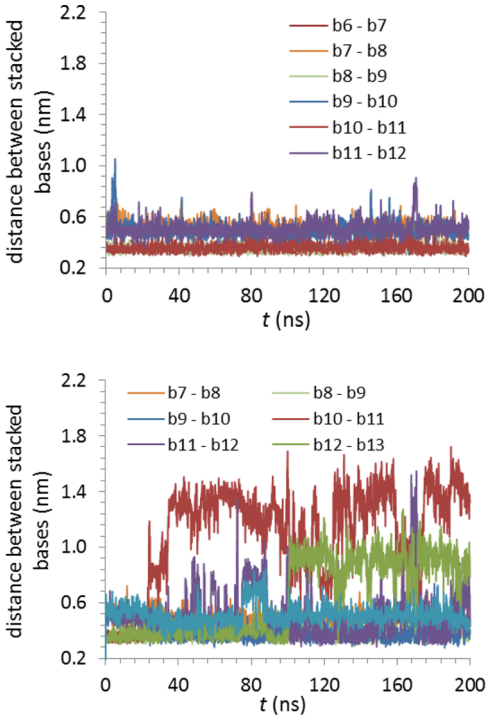


Figure 5. Distances between stacked bases in the siRNA strand I (see text and Table SI) upon LP binding, for cKK-E12-A (top) and cKK-O12 (bottom), as a function of time. A similar graph for the second siRNA strand II appears in Figure S15.

results for the second strand, strand II in Table SI, are shown in Figure S15). This phenomenon is in accord with the observed sharp increase in the number of close contacts between cKK-O12 and siRNA for the ratio 1:1, and also, with the increase of the R_g observed in the siRNA (cf. Figures 4 and 5, and Figure S9).

Moreover, as shown in Figure 6, the distance between opposing bases in some of the base pairs increases considerably

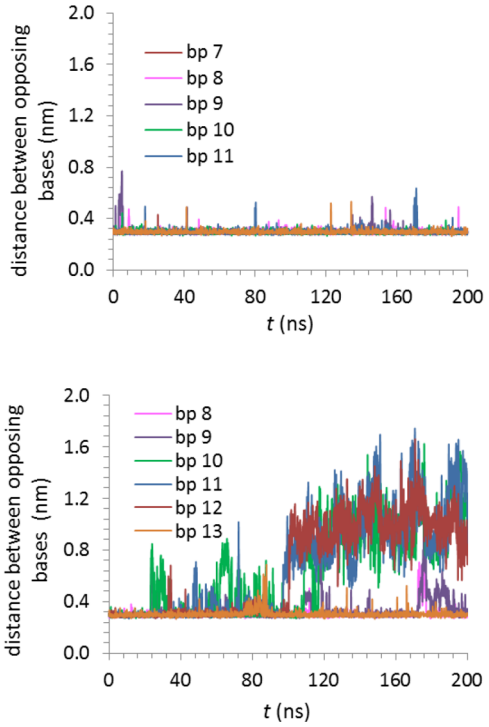


Figure 6. Distances between opposing bases in the siRNA base pairs upon LP binding (see text and Table SI), for cKK-E12-A (top), and cKK-O12 (bottom), as a function of time.

upon the cKK-O12 complexation, reflecting disruption of the Watson−Crick base pairing; this leads to base pair opening and extrusion of bases from the interior of the double helix. By contrast, as shown in Figures 4, 5, and 6, cKK-E12-A does not distort the siRNA strands, thus forming a stable complex.

The extrusion of bases from the interior of the double helix, a motion known as "base flipping",[77,78] is enabled by the rotation of the nucleic acid backbone bonds and occurs in this study because of the local disruption of the Watson−Crick base pairing caused by the insertion of cKK-O12 in the double helix, as depicted in the rendered snapshots from the MD trajectory in Figure 7.

It is noteworthy that "base flipping" has been observed in biological processes such as the DNA repair or transcription[78] and is a result of the change of pH in the U6 RNA intramolecular stem loop (ISL) structure, which is one of the five small nuclear RNAs of the spliceosome.[79] In the latter study, nuclear magnetic resonance spectroscopy showed that base flipping in the U6 RNA ISL is not restricted to a local conformational change, but it rather alters the alignment of the upper and lower helices.[79]

These distortions of the siRNA by cKK-O12 may be reversible or not, as shown by the fluctuations of distances in Figures 5 and 6; hence they can give rise to destabilization of the double helix, destroying this way the siRNA's *in vivo* silencing efficacy. This finding may explain the reason that nanoparticles containing LPs synthesized from epoxide derivatives, like cKK-E12, appear more potent *in vivo* than nanoparticles containing LPs synthesized from acrylate derivatives, like cKK-O12.[16]

Intermolecular Hydrogen Bonding. Hydrogen bonds proved to be a crucial parameter for engineering diverse microscopic thermodynamic processes, spanning sorption thermodynamics and transport phenomena in coordination polymers,[80−82] up to ligand−receptor binding dynamics.[83,84] Computations of HB formation in gene carrier systems have

https://dx.doi.org/10.1021/acs.jctc.9b01261
J. Chem. Theory Comput. 2020, 16, 3842−3855

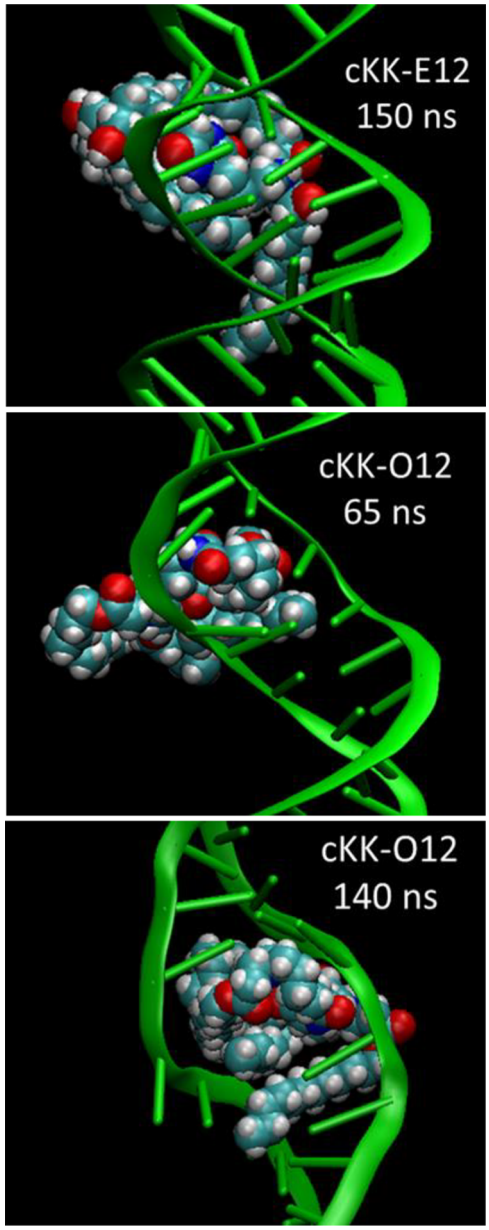


Figure 7. Molecular dynamics snapshots depicting cKK-E12-A and cKK-O12 upon binding to siRNA double helix. The latter LP disrupts the Watson−Crick base pairing. Color code: nitrogen (blue), oxygen (red), carbon (cyan), hydrogen (white). siRNA backbone and bases are depicted as green rods; water and ions are omitted.

been performed in dendrimer−siRNA simulations;[19] nonetheless, the focus of those works has been on HBs either between the siRNA and dendrimer or within the dendrimer. In this study, we investigated the hydrogen bond formation for LP−siRNA, LP−water, and LP−LP.

In Figure 8 the total number of HBs per LP formed with siRNA and water molecules is presented for the highest LP concentration, computed over the entire MD trajectory. The average number of HBs of the hydrated LP in the absence of siRNA is also shown for comparison. The LPs form HBs with both water and siRNA, giving rise to a hydrogen bonding antagonistic scheme; cKK-E12-A forms more HBs with water, compared to results for cKK-O12. The computed HB length and angle distributions (Figure S16) show that HBs between cKK-O12 and siRNA are stronger (vs cKK-O12 with water), whereas

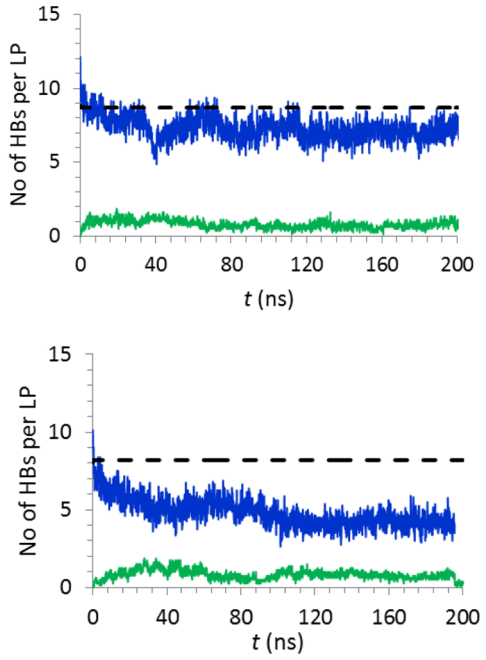


Figure 8. Time evolution of the number of HBs per LP that cKK-E12-A (top) and cKK-O12 (bottom) form with siRNA (green) and water molecules (blue), at the 8:1 LP:siRNA ratio. Lines depict the average LP−water hydrogen bonds in the absence of siRNA for the sake of comparison.

HBs between cKK-E12-A and water are stronger (vs cKK-E12-A with siRNA). Also, the HBs formed by cKK-E12-A with water are stronger (vs cKK-O12). Similar results are obtained for the lower LP concentrations (Figures S17 and S18). The LP−LP hydrogen bonding was also explored in the multi-LP systems; computations showed that no HBs were formed among the LPs in the 4:1 and 8:1 LP:siRNA systems.

The above identified multiple and stronger HBs of cKK-E12-A (vs cKK-O12) with water can be attributed to the hydroxyl FGs located on its tails—a finding consistent with the identified enhanced hydrophilicity of cKK-E12 vs cKK-O12 in aqueous solution, discussed in section A. Moreover, the FGs are located at positions on the hydrocarbon tails that favor the HB pairing with water on a steric basis, as opposed to siRNA. The reader is reminded that the FGs located close to the neutral tertiary amine group reside on average further away from siRNA, thus favoring HB pairings with water.

The weaker HBs between cKK-O12 and water (vs cKK-E12-A), combined with the inherently higher structural flexibility of its longer hydrophobic tails because of the ester linkages (see section A), may provide a microscopic explanation for the disruption of the base pairing in the siRNA structure and its insertion in the double helix, as discussed in the preceding section.

Our results show that lipopeptides participate in an intermolecular hydrogen bonding tug-of-war effect, thus suggesting that chemical modification through the insertion of functional groups of different polarities in proper positions on their tails influences the dynamics of the complex formation. In a more specific context, the incorporation of polar groups in the carrier's structure at positions that favor sterically hydrogen bond pairings with the surrounding solvent could deploy ways toward modulating the dynamics of gene carrier supramolecular assembly.

https://dx.doi.org/10.1021/acs.jctc.9b01261
J. Chem. Theory Comput. 2020, 16, 3842−3855

C. Effect of Stereoisomerism. The effect of the lipopeptides' stereoisomerism on siRNA complexation is investigated in this section by modeling the two cKK-E12 diastereomers as described in the preceding computer reconstruction section. Molecular dynamics simulations of the protonated diastereomers in the presence of siRNA in water and ions were conducted at the temperature of 310 K and pressure of 1 bar, for the time length of 200 ns. Computations were also carried out at 290 and 300 K, and then postprocessing of the dynamics trajectories was performed following the same procedure as previously presented.

It is noteworthy that, although both diastereomers have similar radii of gyration in aqueous solution (see section A), the R_g of cKK-E12-A increases further when bound to siRNA (Figure S19). Also, cKK-E12-A forms a higher number of close contacts with siRNA, as shown by the NoCC presented in Figure 9. These results indicate that cKK-E12-A achieves more open conformations upon binding, exposing a higher number of interaction sites to siRNA.

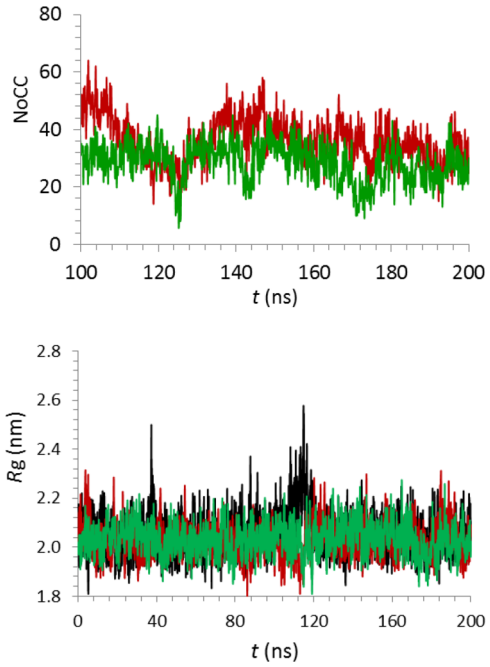


Figure 9. Top: variation of the number of close contacts (defined here as the number of LP atoms that are closer than 0.3 nm from every atom of the siRNA) during the last 100 ns of an equilibrated MD trajectory, for the diastereomers: cKK-E12-A (red) and cKK-E12-B (green). Bottom: time evolution of the siRNA radius of gyration upon binding with cKK-E12-A (red) and cKK-E12-B (green). The R_g of the naked siRNA in aqueous solution (black) is also presented for the sake of comparison.

The differences mentioned above can be attributed to steric effects due to the different spatial arrangements of the hydrocarbon branches and tails of the lipopeptide diastereomers (see Materials Modeling section) that produce dissimilar configurations upon interacting with the chiral oligonucleotide. Namely, the functional groups of cKK-E12-B, FG1 and FG2, which are located close to protonated SJ1, reside on average further away from siRNA, compared with the respective FGs of cKK-E12-A that bind closely to siRNA (Figure S20). In addition, contrary to cKK-E12-A, the terminal methyl groups of cKK-E12-B retain on average longer distances from siRNA and exhibit larger fluctuations (Figure S20), thus indicating more mobile tails for cKK-E12-B. It is notable that both diastereomers form practically equivalent HBs in number and strength with water (Figures S21).

The above findings reveal that LP diastereomers exhibit different binding dynamics upon complexation. Also, they do not induce any conformational change to the siRNA macromolecule, as can be seen in Figure 9 from the time evolution of the computed radius of gyration.

D. Binding Thermodynamics. We calculated the free energy of binding of the complexed system, $[\mathrm{LP}\cdots\mathrm{siRNA}]_{aq}$ described in the Materials Modeling section, by conducting MD simulations at different temperatures in order to estimate the entropic and enthalpic part of the supramolecular process as well. For the postprocessing of the produced trajectories, the aforementioned eq 6 was utilized in the form of eq 9, which allows both contributions to total potential, van der Waals and Coulombic, to vary accordingly.[85]

$$\Delta_b G = \alpha\Delta\langle\mathcal{V}_d\rangle + \beta\Delta\langle\mathcal{V}_e\rangle + \gamma \quad (9)$$

The van der Waals type interactions were modeled here via potential functions embodying nonbonded interactions between skeletal atoms of the biomolecules, as well as interactions depending on the deviations of bond lengths (stretching), bond angles (bending), and dihedral angles (torsion) from their equilibrium values. The electrostatic part was represented by Coulombic interactions between partial charges on atoms. Equation 9, thus reflects the dependence of the macromolecules' atoms collision diameters (i.e., Lennard-Jones sigmas), on their partial charges (ions' strength, hence size). The parameter, γ, was introduced in eq 9 to account for the energy cost due to solute cavity formation.[85]

Calibration of parameters α, β, and γ has been reported in the literature through fitting the estimated $\Delta_b G$ to experimental values in a wide group of ligand−protein systems depending on the charge of the ligand or the number of OH groups present on it.[71,86,87] In a recent study on solvation energies for several organic compounds, parametrization of eq 9 was carried out to take account of the nonhydrocarbon groups in the solute structure, e.g., ions, hydroxyls, amides, amines, esters, and carboxyl groups; also, an expression was devised to estimate, β, by means of weighting coefficients for each nonhydrocarbon chemical moiety in the solute.[88]

In the absence of experimental data for the lipopeptide−nucleic acid systems, we computed $\Delta_b G$, by employing the various parametrization schemes mentioned above, and presented in Table 1, as follows: set 1, using a uniform set of parameters;[71] set 2, using parameters to account for the presence of OH groups in the LP structure;[86,87] set 3, using values that involve parameters considering the charge (protonation) of the LP;[86] and set 4, using parametrization that takes all the LP nonhydrocarbon chemical moieties into account by employing values calibrated for organic compounds in solute−solvent systems.[88] It is noted that in all parametrization schemes, the γ value was kept to zero;[71,86,88] however, parametrization of eq 9 assigning a nonzero value to the γ parameter did not show any improvement in the calculation of ligand−protein free energies of binding compared to experimental results.[86]

In Figure 10 the presented functions $\Delta_b G(T)$ for the three LPs of this study were computed using the parameters from set 3 (graphs including other sets are shown in Figure S22). The ensemble averages of the potential energies were recorded over successive equilibrated MD trajectories, along with the

https://dx.doi.org/10.1021/acs.jctc.9b01261
J. Chem. Theory Comput. 2020, 16, 3842−3855

Table 1. Thermodynamic Properties of the $[LP\cdots siRNA]_{aq}$ Complexed System at $T = 310$ K and $P = 1$ bar, Calculated from MD, for Various Parametrization Sets (See Text) of Eq 9

		parameters $\gamma = 0$				
LP	set	α	β	$\Delta_b G$ (kJ mol^{-1})	$-T\Delta_b S$ (kJ mol^{-1})	$\Delta_b H$ (kJ mol^{-1})
cKK-O12	1	0.161	0.5	−41.1 ± 20.7	−255.75	214.6
	2	0.181	0.43	−45.3 ± 18.0	−232.5	187.2
	3	0.181	0.5	−45.8 ± 20.8	−262.0	216.2
	4	0.181	0.482	−45.6 ± 20.0	−252.6	207.0
cKK-E12-A	1	0.161	0.5	−50.0 ± 21.2	−316.2	266.2
	2	0.181	0.33	−42.6 ± 13.7	−184.4	141.8
	3	0.181	0.5	−52.5 ± 21.0	−308.5	256.0
	4	0.181	0.47	−50.8 ± 19.7	−286.7	235.9
cKK-E12-B	1	0.161	0.5	−38.1 ± 23.4	−389.0	351.0
	2	0.181	0.33	−30.6 ± 15.2	−192.2	161.6
	3	0.181	0.5	−39.6 ± 23.3	−372.0	332.4
	4	0.181	0.47	−38.0 ± 21.8	−341.0	303.0

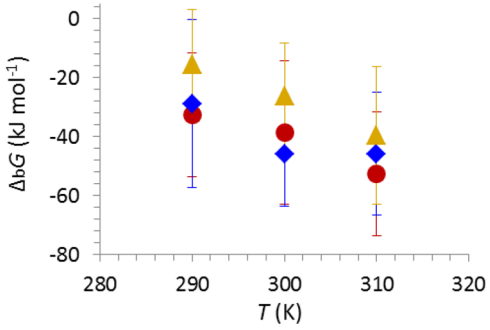


Figure 10. Free energies of binding for the complexed system: $[LP\cdots siRNA]_{aq}$ as a function of temperature, for the LPs: cKK-O12 (rhombi) and the diastereomers cKK-E12-A (circles) and cKK-E12-B (triangles).

corresponding error bars evaluated from the standard deviations of the electrostatic and van der Waals type parts that constitute the potential energy of each system.

In view of the trend of the predicted values one might safely interpret the depicted free energy functions as linearly dependent on temperature, to obtain a first estimate of entropy via the partial derivative of eq 7. In fact, this hypothesis, in view of the derivative of eq 8, would mean that the heat capacities of the two systems at the given temperature range may be considered nearly equal; in other words, the existence of siRNA does not alter significantly the heat capacity of the aqueous LP system 1, resulting in a practically temperature independent entropy of binding within the examined temperature range. The latter hypothesis, supported by the simulated data within the narrow temperature regime, is not expected to have a discernible effect otherwise.

As seen in Table 1, the computed free energies of binding for siRNA with cKK-O12 and the diastereomers cKK-E12-A and cKK-E12-B are all negative for all the parameter sets tested, indicating indeed a natural (thermodynamically irreversible) complex formation process, wherein the Gibbs energy function under the prescribed constant temperature and pressure must take a minimum value.

The average energy difference term, $P\Delta\langle\overline{V}\rangle$, between systems 1 and 2, where $\overline{V}$ denotes intensive system volume, vanishes (it remained almost unaltered between the two systems; see the Supporting Information); thus, the positive enthalpic contribution to $\Delta_b G$ reflects the unfavorable atomic interactions in the complexation scheme under study. Therefore, $\Delta_b H$ can be considered approximately equal to $\Delta_b U$, the molar configurational internal energy of binding, that is to say, the computed potential energy difference. Thus, $\Delta_b H$ manifests the strength of interacting sites and/or number of close contacts, plus the energy cost due to water displacement by the LPs. Furthermore, the positive entropy of binding, $\Delta_b S$, reflects the complexed system's total gain in configurational degrees of freedom, plus the contribution from the aforesaid water freeing up, and suggests that this complexation process is predominantly entropy driven.

The computed negative correlation between enthalpy and the entropic term, $-T\Delta_b S$, is usually observed in ligand−protein systems.[89] Entropy driven binding has also been observed in the isothermal titration calorimetry experiments of alkylammonium lipids[90] and lipospermine[91] when binding to DNA. These lipid-like molecules have structures comprising a polar ammonium headgroup and single or multiple long aliphatic chains akin to the LPs' structure of this study; the experimental values of the Gibbs energy of binding to DNA for dodecylammonium[90] and lipospermine[91] molecules, $\Delta_b G = -22.9$ kJ mol^{-1} and $\Delta_b G = -37.5$ kJ mol^{-1} at 298 K, respectively, are in good agreement with our computed values for the LP−siRNA complexes (see Table 1). In the above lipid experiments, positive enthalpies of binding were related to the formation of a mobile and disordered lipid coating on the surface of DNA, based on the "assumption" that a negative enthalpic part, measured upon aggregation of the same lipids from aqueous solution into a pure lipid solid phase,[92] would reflect the formation of a solid lipid phase.[90]

The above assumption is in fact a thermodynamics fundamental for such transitions (dense to denser phase); however, enthalpy itself is by no means the proper proxy for free energy (the true potential of the process). In our work this is validated by an endothermic process, which through the formation of a less ordered phase, points to a negative binding free energy of the complexed system $[LP\cdots siRNA]_{aq}$, being in accord with the analysis of dynamics and critical sites in section B.

As seen in Table 1, the diastereomers cKK-E12-A and cKK-E12-B, regardless of the parameter set used, present different binding energies. The lower $\Delta_b G$ for cKK-E12-A, indicates a thermodynamically more stable complex. The cKK-E12-B diastereomer exhibits higher tail mobility, and fewer close contacts, with the siRNA than cKK-E12-A does (see section C); these features are in agreement with the computed higher

https://dx.doi.org/10.1021/acs.jctc.9b01261
J. Chem. Theory Comput. 2020, 16, 3842−3855

configurational entropy gain and higher potential energy, respectively, for cKK-E12-B when bound to siRNA.

It should be stressed that a complex of low binding energy does not necessarily point to a more efficient carrier; more importantly, the capability of a nanoparticle to release the siRNA to the cytosol is one of the critical design characteristics toward successful gene carriers. In other words, one must be able to control the interplay between the biomolecular complex stability and its susceptibility to dissociation. The looser binding of the diastereomer cKK-E12-B may facilitate the siRNA escape in the cytosol at the final stage of the delivery process, which is a prerequisite for the proper function of the gene via the RNAi pathway.

Effect of Salt Concentration. We investigated the effect of salt on the LP:siRNA complexation by performing molecular dynamics simulations and free energy computations for the 1:1 cKK-E12-A:siRNA system at physiological salt concentration, namely, 150 mM NaCl, at 310 K and pressure of 1 bar. The conformational stability of the siRNA double helix after the complexation with cKK-E12-A at the above salt concentration is preserved, since no disruption of the Watson−Crick base pairing is observed, as depicted by the computed distances between opposing bases in the siRNA base pairs in Figure S23.

We also found that the binding energy of cKK-E12-A to siRNA increases, namely, from −52.5 to −26.9 kJ mol^{-1} (using parametrization set 3), at physiological salt concentration. The computed $\Delta_b G$ values, using the parametrization sets described in the previous paragraph, can be seen in Table SV. Our findings are consistent with the $\Delta_b G$ values of dodecylammonium−DNA binding upon an increase in salt concentration, namely, from −22.9 to −16.0 kJ mol^{-1} at 10 and 200 mM NaCl respectively, as reported in isothermal titration calorimetry experiments.[90]

CONCLUSIONS

Molecular dynamics simulations in atomistic detail coupled with free energy computations were performed to explore the binding dynamics and thermodynamics of siRNA with selected epoxide- and acrylate-derived lipopeptides, members of a recently synthesized class of lipopeptide structures that constitute the key component of lipid-based nanocarriers intended as potent RNAi therapeutics.

We found that both LP types bind preferentially to the minor groove of the siRNA, primarily with their protonated tertiary amine group localized in close proximity to the siRNA electronegative atoms, while their lipid-like tails exhibit labile contacts with the siRNA, thus forming a mobile disordered coating, which surrounds the biomacromolecule.

The acrylate-derived LP, coded as cKK-O12, binds to siRNA and distorts the macromolecule's structure by disrupting the Watson−Crick base pairing and inducing a base flipping motion. By contrast, the epoxide-derived LP, coded as cKK-E12, does not induce conformational changes to siRNA, which thus remains intact during the simulation timelength. This finding can explain the observed higher *in vivo* potency of nanoparticles containing LPs synthesized from epoxide derivatives in comparison to nanoparticles containing LPs synthesized from acrylate derivatives.

These differences can be explained on the basis of the type of functional groups which are present in the LP tails. The presence of hydroxyl groups at the tails of cKK-E12 results in stronger HB pairings with water than the siRNA. The ester groups of the cKK-O12 tails, however, exhibit weaker HBs with water. The latter bonding effect, combined with the inherently higher flexibility of the cKK-O12 tails due to the presence of ester linkages are responsible for the observed more compact conformations, which eventually result in disrupting the siRNA double helix.

This competing hydrogen bond pairing of the LP−water versus LP−siRNA resembles a HB tug-of-war effect and reveals the water's improved role apart from being simply a passive solvent surrounding the complex. It also underpins a significant molecular design parameter by the incorporation of polar functionalities in proper positions of the carrier's structure that favor the HB pairing with water, providing the synthesis people with possible routes toward modulating the association dynamics of the self-assembled supramolecular complex.

The computational binding thermodynamics shows a negative entropic contribution, $-T\Delta_b S$, due to increasing disorder upon complexation ($\Delta_b S > 0$), as opposed to a positive enthalpic part, suggesting that the binding of lipopeptides to siRNA is an endothermic process driven predominantly by entropy. Interestingly, contrary to the belief that drugs (in general, ligands), which bind due to negative, hence favorable, enthalpies have certain advantages over ligands where binding is merely driven by entropy (because enhanced interactions improve selective capture, resulting in increased efficacy due to their geometric specificity), our findings concerning the predominance of configurational entropy in gene-based drugs wherein molecular size and flexibility are major issues, must not be surprising. Additional computations show that the binding energy increases at increased (physiological) salt concentration, consistent with isothermal titration calorimetry experiments reported in similar systems.

Because the synthesis reactions of these lipopeptides are not stereoselective, a mixture of diastereoisomers is obtained. By examining two diastereomers of the cKK-E12 lipopeptide, we found that both of them form stable complexes without distorting the siRNA double helix. Nevertheless, they exhibit structural and energetics differences upon binding to siRNA due to steric effects imposed by the different absolute configurations (*R* or *S*) at their chiral centers. Thus, although the diastereomer cKK-E12-A forms a thermodynamically stronger complex, the relatively looser binding of the diastereomer cKK-E12-B may facilitate the siRNA escape in the cytosol at the final stage of the delivery process.

Thus, stereoisomerism can serve as an efficient design parameter toward modulating the strength of the complex, encouraging further exploration of stereochemically pure carriers or mixtures of selected diastereomers that may help tuning biomolecular complexes to an optimum strength in order to control the interplay between complex stability and ease of releasing the nucleic acid drugs into the cell.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.9b01261.

Simulation details, hydrogen bond notations and computational details, figures of structural formulas, time evolution graphs, solution configurations, bond and angle distributions, MD trajectory snapshots, average minimum distance vs nucleotide pair, distances between bases vs time, $\Delta_b G$ vs temperature, and tables of sequences of bases, force field atom types, time averages and standard deviations of R_g, and $\Delta_b G$ values (PDF)

https://dx.doi.org/10.1021/acs.jctc.9b01261
J. Chem. Theory Comput. 2020, 16, 3842−3855

AUTHOR INFORMATION

Corresponding Authors

Evangelia Pantatosaki − *School of Chemical Engineering, National Technical University of Athens, 15780 Athens, Greece;* orcid.org/0000-0002-2063-097X; Email: evpanta@chemeng.ntua.gr

George K. Papadopoulos − *School of Chemical Engineering, National Technical University of Athens, 15780 Athens, Greece; Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States;* orcid.org/0000-0001-5935-5461; Email: gkpap@chemeng.ntua.gr, georgekp@mit.edu

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.9b01261

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

E.P. acknowledges a grant of Scientific Excellence by the Hellenic Foundation for Research and Innovation (Project: B.1003-ENGETACT). G.K.P. acknowledges support by European Union through a fellowship under H2020-MSC-IF (Project: 796794-ENGEMED) and by the Greek Research and Technology Network for CPU-time provided in the National HPC facility "ARIS".

REFERENCES

(1) Belhke, M. A. Progress towards *in vivo* use of siRNAs. *Mol. Ther.* **2006**, *13* (4), 644−670.

(2) Scherman, D.; Rousseau, A.; Bigey, P.; Escriou, V. Genetic pharmacology: progresses in siRNA delivery and therapeutic applications. *Gene Ther.* **2017**, *24*, 151−156.

(3) Fire, A.; Xu, S. Q.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E.; Mello, C. C. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **1998**, *391*, 806−811.

(4) Turner, J. J.; Jones, S. W.; Moschos, S. A.; Lindsay, M. A.; Gait, M. J. MALDI-TOF mass spectral analysis of siRNA degradation in serum confirms an RNAse A-like activity. *Mol. BioSyst.* **2007**, *3*, 43−50.

(5) Marques, J. T.; Williams, B. R. Activation of the mammalian immune system by siRNAs. *Nat. Biotechnol.* **2005**, *23*, 1399−1405.

(6) Behlke, M. A. Chemical Modification of siRNAs for In Vivo Use. *Oligonucleotides* **2008**, *18*, 305−320.

(7) Thomas, C. E.; Ehrhardt, A.; Kay, M. A. Progress and problems with the use of viral vectors for gene therapy. *Nat. Rev. Genet.* **2003**, *4* (5), 346−358.

(8) Peer, D.; Lieberman, J. Special delivery: targeted therapy with small RNAs. *Gene Ther.* **2011**, *18*, 1127−1133.

(9) Kanasty, R.; Dorkin, J. R.; Vegas, A.; Anderson, D. Delivery materials for siRNA therapeutics. *Nat. Mater.* **2013**, *12*, 967−977.

(10) Tam, Y. C.; Chen, S.; Cullis, P. R. Advances in Lipid Nanoparticles for siRNA Delivery. *Pharmaceutics* **2013**, *5*, 498−507.

(11) Lorenzer, C.; Dirin, M.; Winkler, A.-M.; Baumann, V.; Winkler, J. Going beyond the liver: progress and challenges of targeted delivery of siRNA therapeutics. *J. Controlled Release* **2015**, *203*, 1−15.

(12) Huang, Y.; Hong, J.; Zheng, S.; Ding, Y.; Guo, S.; Zhang, H.; Zhang, X.; Du, Q.; Liang, Z. Elimination Pathways of Systemically Delivered siRNA. *Mol. Ther.* **2011**, *19* (2), 381−385.

(13) Ledford, H. Gene-silencing drug approved. *Nature* **2018**, *560*, 291−292.

(14) Wan, C.; Allen, T. M.; Cullis, P. R. Lipid nanoparticle delivery systems for siRNA-based therapeutics. *Drug Delivery Transl. Res.* **2014**, *4* (1), 74−83.

(15) Love, K. T.; Mahon, K. P.; Levins, C. G.; Whitehead, K. A.; Querbes, W.; Dorkin, J. R.; Qin, J.; Cantley, W.; Qin, L. L.; Racie, T.; Frank-Kamenetsky, M.; Yip, K. N.; Alvarez, R.; Sah, D. W. Y.; de Fougerolles, A.; Fitzgerald, K.; Koteliansky, V.; Akinc, A.; Langer, R.; Anderson, D. G. Lipid-like materials for low-dose, in vivo gene silencing. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 1864−1869.

(16) Dong, Y.; Love, K. T.; Dorkin, J. R.; Sirirungruang, S.; Zhang, Y.; Chen, D.; Bogorad, R. L.; Yin, H.; Chen, Y.; Vegas, A. J.; Alabi, C. A.; Sahay, G.; Olejnik, K. T.; Wang, W.; Schroeder, A.; Lytton-Jean, A. K. R.; Siegwart, D. J.; Akinc, A.; Barnes, C.; Barros, S. A.; Carioto, M.; Fitzgerald, K.; Hettinger, J.; Kumar, V.; Novobrantseva, T. I.; Qin, J.; Querbes, W.; Koteliansky, V.; Langer, R.; Anderson, D. G. Lipopeptide nanoparticles for potent and selective siRNA delivery in rodents and nonhuman primates. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 3955−3960.

(17) Vasumathi, V.; Maiti, P. K. Complexation of siRNA with Dendrimer: A Molecular Modeling Approach. *Macromolecules* **2010**, *43*, 8264−8274.

(18) Ouyang, D.; Zhang, H.; Herten, D.-P.; Parekh, H. S.; Smith, S. C. Structure, Dynamics, and Energetics of siRNA-Cationic Vector Complexation: A Molecular Dynamics Study. *J. Phys. Chem. B* **2010**, *114*, 9220−9230.

(19) Karatasos, K.; Posocco, P.; Laurini, E.; Pricl, S. Poly-(amidoamine)-based Dendrimer/siRNA Complexation Studied by Computer Simulations: Effects of pH and Generation on Dendrimer Structure and siRNA Binding. *Macromol. Biosci.* **2012**, *12*, 225−240.

(20) Pavan, G. M.; Mintzer, M. A.; Simanek, E. E.; Merkel, O. M.; Kissel, T.; Danani, A. Computational Insights into the Interactions between DNA and siRNA with "Rigid" and "Flexible" Triazine Dendrimers. *Biomacromolecules* **2010**, *11*, 721−730.

(21) Pavan, G. M.; Kostiainen, M. A.; Danani, A. Computational Approach for Understanding the Interactions of UV-Degradable Dendrons with DNA and siRNA. *J. Phys. Chem. B* **2010**, *114*, 5686−5693.

(22) Sun, C.; Tang, T.; Uludag, H. A molecular dynamics simulation study on the effect of lipid substitution on polyethylenimine mediated siRNA complexation. *Biomaterials* **2013**, *34*, 2822−2833.

(23) Ziebarth, J. D.; Kennetz, D. R.; Walker, N. J.; Wang, Y. Structural Comparisons of PEI/DNA and PEI/siRNA Complexes Revealed with Molecular Dynamics Simulations. *J. Phys. Chem. B* **2017**, *121*, 1941−1952.

(24) Lu, Y.; Li, J.; Lu, D. The mechanism for the complexation and dissociation between siRNA and PMAL: a molecular dynamics simulation study based on a coarse-grained model. *Mol. Simul.* **2017**, *43* (13−16), 1385−1393.

(25) Kim, M.; Kim, H. R.; Chae, S. Y.; Larson, R. G.; Lee, H.; Park, J. C. Effect of Arginine-Rich Peptide Length on the Structure and Binding Strength of siRNA-Peptide Complexes. *J. Phys. Chem. B* **2013**, *117*, 6917−6926.

(26) McConathy, J.; Owens, M. J. Stereochemistry in Drug Action. *Prim Care Companion J. Clin Psychiatry* **2003**, *5* (2), 70−73.

(27) Chhabra, N.; Aseri, M. L.; Padmanabhan, D. A review of drug isomerism and its significance. *Int. J. Appl. Basic Med. Res.* **2013**, *3* (1), 16−18.

(28) Schultze, P.; Feigon, J. Chirality errors in nucleic acid structures. *Nature* **1997**, *387*, 668.

(29) Herman, D. M.; Baird, E. E.; Dervan, P. B. Stereochemical Control of the DNA Binding Affinity, Sequence Specificity, and Orientation Preference of Chiral Hairpin Polyamides in the Minor Groove. *J. Am. Chem. Soc.* **1998**, *120*, 1382−1391.

(30) Arjmand, F.; Sayeed, F.; Parveen, S.; Tabassum, S.; Juvekar, A. S.; Zingde, S. M. Design and synthesis of (S)- and (R)-enantiomers of [4-(2-hydroxy-1-phenylethylimino)pent-2-ol]-dimethyltin(IV) and 2,2-dimethyl-4-phenyl-1,3,2-oxazastannolidine: in vitro antitumor activity against human tumor cell lines and in vivo assay of (S)-enantiomers. *Dalton Trans.* **2013**, *42*, 3390−3401.

(31) (a) Cahn, R. S.; Ingold, C.; Prelog, V. Specification of Molecular Chirality. *Angew. Chem., Int. Ed. Engl.* **1966**, *5*, 385−415. (b) Cross, L. C.; Klyne, W. Rules for the Nomenclature of Organic Chemistry, Section E: Stereochemistry. *Pure Appl. Chem.* **1976**, *45*, 11−30.

(32) Schroeder, A.; Levins, C. G.; Cortez, C.; Langer, R.; Anderson, D. G. Lipid-based nanotherapeutics for siRNA delivery. *J. Intern. Med.* **2010**, *267*, 9−21.
(33) Yuan, Y. R.; Pei, Y.; Chen, H. Y.; Tuschl, T.; Patel, D. J. A Potential Protein-RNA Recognition Event along the RISC-Loading Pathway from the Structure of A. aeolicus Argonaute with Externally Bound siRNA. *Structure* **2006**, *14*, 1557−1565.
(34) (a) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1−2*, 19−25. (b) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible and Free. *J. Comput. Chem.* **2005**, *26*, 1701−1718. (c) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **1995**, *91*, 43−56.
(35) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dorr, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the AMBER ff99SB protein force field. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 1950−1958.
(36) Balaceanu, A.; Pérez, A.; Dans, P. D.; Orozco, M. Allosterism and signal transfer in DNA. *Nucleic Acids Res.* **2018**, *46*, 7554−7565.
(37) Marklund, E. G.; Mahmutovic, A.; Berg, O. G.; Hammar, P.; van der Spoel, D.; Fange, D.; Elf, J. Transcription-factor binding and sliding on DNA studied using micro- and macroscopic models. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 19796−19801.
(38) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(39) Kaminski, G.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.
(40) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(41) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E., III; Cruzeiro, V. W. D.; et al. *AMBER 2018*; University of California: San Francisco, 2018.
(42) Cheatham, T. E., 3rd; Cieplak, P.; Kollman, P. A. A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dyn.* **1999**, *16*, 845−862.
(43) Wang, J.; Cieplak, P.; Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **2000**, *21*, 1049−1074.
(44) Šponer, J.; Cang, X.; Cheatham, T. E. Molecular dynamics simulations of G-DNA and perspectives on the simulation of nucleic acid structures. *Methods* **2012**, *57*, 25−39.
(45) Vangaveti, S.; Ranganathan, S. V.; Chen, A. A. Advances in RNA molecular dynamics: a simulator's guide to RNA force fields. *WIREs RNA* **2017**, *8*, e1396.
(46) Šponer, J.; Bussi, G.; Krepl, M.; Banáš, P.; Bottaro, S.; Cunha, R. A.; Gil-Ley, A.; Pinamonti, G.; Poblete, S.; Jurečka, P.; Walter, N. G.; Otyepka, M. RNA Structural Dynamics As Captured by Molecular Simulations: A Comprehensive Overview. *Chem. Rev.* **2018**, *118*, 4177−4338.
(47) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. Refinenement of the AMBER Force Field for Nucleic Acids: Improving the Description of Alpha/Gamma Conformers. *Biophys. J.* **2007**, *92*, 3817−3829.
(48) Zgarbová, M.; Otyepka, M.; Šponer, J.; Mládek, A.; Banáš, P.; Cheatham, T. E.; Jurečka, P. Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J. Chem. Theory Comput.* **2011**, *7*, 2886−2902.
(49) Banáš, P.; Hollas, D.; Zgarbová, M.; Jurečka, P.; Orozco, M.; Cheatham, T. E., III; Šponer, J.; Otyepka, M. Performance of Molecular Mechanics Force Fields for RNA Simulations: Stability of UUCG and GNRA Hairpins. *J. Chem. Theory Comput.* **2010**, *6*, 3836−3849.
(50) Tan, D.; Piana, S.; Dirks, R. M.; Shaw, D. E. RNA force field with accuracy comparable to state-of-the-art protein force fields. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, E1346−E1355.
(51) Piana, S.; Donchev, A. G.; Robustelli, P.; Shaw, D. E. Water dispersion interactions strongly influence simulated structural properties of disordered protein states. *J. Phys. Chem. B* **2015**, *119*, 5113−5123.
(52) Kührová, P.; Mlýnsky′, V.; Zgarbová, M.; Krepl, M.; Bussi, G.; Best, R. B.; Otyepka, M.; Šponer, J.; Banáš, P. Improving the Performance of the Amber RNA Force Field by Tuning the Hydrogen-Bonding Interactions. *J. Chem. Theory Comput.* **2019**, *15*, 3288−3305.
(53) Robertson, M. J.; Qian, Y.; Robinson, M. C.; Tirado-Rives, J.; Jorgensen, W. L. Development and Testing of the OPLS-AA/M Force Field for RNA. *J. Chem. Theory Comput.* **2019**, *15*, 2734−2742.
(54) *GROMACS* 2020.1, Reference Manual, accessed on 14 April 2020, http://manual.gromacs.org/documentation/current/reference-manual/topologies/force-field-organization.html.
(55) Dans, P. D.; Gallego, D.; Balaceanu, A.; Darré, L.; Gómez, H.; Orozco, M. Modeling, Simulations, and Bioinformatics at the Service of RNA Structure. *Chem.* **2019**, *5*, 51−73.
(56) Sklenovsky, P.; Florova, P.; Banas, P.; Reblova, K.; Lankas, F.; Otyepka, M.; Sponer, J. Understanding RNA Flexibility Using Explicit Solvent Simulations: The Ribosomal and Group I Intron Reverse Kink-Turn Motifs. *J. Chem. Theory Comput.* **2011**, *7*, 2963−2980.
(57) Kührová, P.; Otyepka, M.; Šponer, J.; Banáš, P. Are Waters around RNA More than Just a Solvent? − An Insight from Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 401−411.
(58) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(59) Galindo-Murillo, R.; Bergonzo, C.; Cheatham, T. E., III Molecular Modeling of Nucleic Acid Structure: Setup and Analysis. *Curr. Protoc. Nucleic Acid Chem.* **2014**, *56*, 7.10.1−7.10.21.
(60) Chaplin, M. Water Structure and Science. http://www1.lsbu.ac.uk/water/nucleic_acid_hydration.html#hydr.
(61) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52*, 7182−7190.
(62) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
(63) (a) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463−1472. (b) Hess, B. P-LINCS: A parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 116−122.
(64) (a) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092. (b) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh ewald potential. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(65) Humphrey, W.; Dalke, A.; Schulten, K. VMD - Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.
(66) (a) Singer, S. J.; Chandler, D. Free energy functions in the extended RISM approximation. *Mol. Phys.* **1985**, *55*, 621−625. (b) Carlson, H. A.; Jorgensen, W. L. An Extended Linear Response Method for Determining Free Energies of Hydration. *J. Phys. Chem.* **1995**, *99*, 10667−10673.
(67) Chandler, D. Equilibrium theory of polyatomic fluids. In *Studies in Statistical Mechanics*; Lebowitz, J. L., Montroll, E. W., Eds.; North Holland: Amsterdam, 1982; Vol, *8*, pp 275−340.
(68) Roux, B.; Yu, H.-A.; Karplus, M. Molecular basis for the Born model of ion solvation. *J. Phys. Chem.* **1990**, *94*, 4683−4688.
(69) Born, M. Volumen und Hydratationswiirme der Ionen. *Eur. Phys. J. A* **1920**, *1*, 45−48.

https://dx.doi.org/10.1021/acs.jctc.9b01261
J. Chem. Theory Comput. 2020, 16, 3842−3855

(70) Yu, A.; Karplus, M. A thermodynamic analysis of solvation. *J. Chem. Phys.* **1988**, *89*, 2366−2379.
(71) (a) Åqvist, J.; Medina, C.; Samuelsson, J.-E. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng., Des. Sel.* **1994**, *7*, 385−391. (b) Åqvist, J.; Luzhkov, V. B.; Brandsdal, B. O. Ligand Binding Affinities from MD Simulations. *Acc. Chem. Res.* **2002**, *35*, 358−365.
(72) Kumar, S. K.; Floriano, M. A.; Panagiotopoulos, A. Z. Nanostructure formation and phase separation in surfactant solutions. *Adv. Chem. Eng.* **2001**, *28*, 297−311.
(73) Arunan, E.; Desiraju, G. R.; Klein, R. A.; Sadlej, J.; Scheiner, S.; Alkorta, I.; Clary, D. C.; Crabtree, R. H.; Dannenberg, J. J.; Hobza, P.; Kjaergaard, H. G.; Legon, A. C.; Mennucci, B.; Nesbitt, D. Defining the hydrogen bond: An Account (IUPAC Technical Report). *Pure Appl. Chem.* **2011**, *83*, 1619−1636.
(74) Khalatur, P. G. The flexibility of aliphatic polyester chains. *Polym. Sci. U.S.S.R.* **1979**, *21*, 2416−2423.
(75) Gastinel-Jasso, C. F.; Soltero-Martinez, J. F. A. Introduction: Modifiable Characteristics and Applications. In *Modification of Polymer Properties*; Jasso-Gastinel, C. F., Kenny, J. M., Eds.; Elsevier Inc., 2017; Chapter 1, pp 15−16.
(76) Luscombe, N. M.; Laskowski, R. A.; Thornton, J. M. Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res.* **2001**, *29* (13), 2860−2874.
(77) Klimasauskas, S.; Kumar, S.; Roberts, R. J.; Cheng, X. Hhal methyltransferase flips its target base out of the DNA helix. *Cell* **1994**, *76* (2), 357−369.
(78) Roberts, R. J. On the base flipping. *Cell* **1995**, *82*, 9−12.
(79) Reiter, N. J.; Blad, H.; Abildgaard, F.; Butcher, S. E. Dynamics in the U6 RNA Intramolecular Stem-Loop: A Base Flipping Conformational Change. *Biochemistry* **2004**, *43*, 13739−13747.
(80) Kolokathis, P. D.; Pantatosaki, E.; Papadopoulos, G. K. Atomistic Modeling of Water Thermodynamics and Kinetics Within MIL-100(Fe). *J. Phys. Chem. C* **2015**, *119*, 20074−20084.
(81) Fröhlich, D.; Pantatosaki, E.; Kolokathis, P. D.; Markey, K.; Reinsch, H.; Baumgartner, M.; van der Veen, M. A.; De Vos, D. E.; Stock, N.; Papadopoulos, G. K.; Henninger, S. K.; Janiak, C. Water adsorption behaviour of CAU-10-H: A thorough investigation of its structure-property relationships. *J. Mater. Chem. A* **2016**, *4*, 11859−11869.
(82) Splith, T.; Pantatosaki, E.; Kolokathis, P. D.; Fröhlich, D.; Zhang, K.; Füldner, G.; Chmelik, C.; Jiang, J.; Henninger, S. K.; Stallmach, F.; Papadopoulos, G. K. Molecular Dynamics Phenomena of Water in the Metalorganic Framework MIL-100(Al), as Revealed by Pulsed Field Gradient NMR and Atomistic Simulation. *J. Phys. Chem. C* **2017**, *121*, 18065−18074.
(83) Chen, D.; Oezguen, N.; Urvil, P.; Ferguson, C.; Dann, S. M.; Savidge, T. C. Regulation of protein-ligand binding affinity by hydrogen bond pairing. *Sci. Adv.* **2016**, *2*, e1501240.
(84) Schiebel, J.; Gaspari, R.; Wulsdorf, T.; Ngo, K.; Sohn, K.; Schrader, T. E.; Cavalli, A.; Ostermann, A.; Heine1, A.; Klebe, G. Intriguing role of water in protein-ligand binding studied by neutron crystallography on trypsin complexes. *Nat. Commun.* **2018**, *9*, 3559.
(85) Carlson, H. A.; Jorgensen, W. L. An Extended Linear Response Method for Determining Free Energies of Hydration. *J. Phys. Chem.* **1995**, *99*, 10667−10673.
(86) Hansson, T.; Marelius, J.; Åqvist, J. Ligand binding affinity prediction by linear interaction energy methods. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 27−35.
(87) Åqvist, J.; Hansson, T. On the Validity of Electrostatic Linear Response in Polar Solvents. *J. Phys. Chem.* **1996**, *100*, 9512−9521.
(88) Almlof, M.; Carlsson, J.; Åqvist, J. Improving the accuracy of the linear interaction energy method for solvation free energies. *J. Chem. Theory Comput.* **2007**, *3*, 2162−2175.
(89) Reynolds, C. H.; Holloway, M. K. Thermodynamics of Ligand Binding and Efficiency. *ACS Med. Chem. Lett.* **2011**, *2*, 433−437.
(90) Matulis, D.; Rouzina, I.; Bloomfield, V. A. Thermodynamics of Cationic Lipid Binding to DNA and DNA Condensation: Roles of Electrostatics and Hydrophobicity. *J. Am. Chem. Soc.* **2002**, *124*, 7331−7342.
(91) Patel, M. M.; Anchordoquy, T. J. Contribution of Hydrophobicity to Thermodynamics of Ligand-DNA Binding and DNA Collapse. *Biophys. J.* **2005**, *88* (3), 2089−2103.
(92) Matulis, D.; Bloomfield, V. A. Thermodynamics of the hydrophobic effect. II. Calorimetric measurement of enthalpy, entropy, and heat capacity of aggregation of alkylamines and long aliphatic chains alkylamines. *Biophys. Chem.* **2001**, *93*, 53−65.

https://dx.doi.org/10.1021/acs.jctc.9b01261
J. Chem. Theory Comput. 2020, 16, 3842−3855