Journal of Molecular Graphics and Modelling 84 (2018) 36–42

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Insights into the structural features and stability of peptide nucleic acid with a D-prolyl-2-aminocyclopentane carboxylic acid backbone that binds to DNA and RNA

Nattawee Poomsuk [a], Tirayut Vilaivan [b], Khatcharin Siriwong [a, *]

[a] Materials Chemistry Research Center, Department of Chemistry and Center of Excellence for Innovation in Chemistry, Faculty of Science, Khon Kaen University, Khon Kaen, 40002, Thailand
[b] Organic Synthesis Research Unit, Department of Chemistry, Faculty of Science, Chulalongkorn University, Phayathai Road, Patumwan, Bangkok, 10330, Thailand

ARTICLE INFO

Article history:
Received 5 May 2018
Received in revised form
6 June 2018
Accepted 8 June 2018
Available online 12 June 2018

Keywords:
Peptide nucleic acid
PNA-DNA
PNA-RNA
Binding free energy
Molecular dynamics simulation

ABSTRACT

Peptide nucleic acid (PNA) is a powerful biomolecule with a wide variety of important applications. In this work, the molecular structures and binding affinity of PNA with a D-prolyl-2-aminocyclopentane carboxylic acid backbone (acpcPNA) that binds to both DNA and RNA were studied using molecular dynamics simulations. The simulated structures of acpcPNA-DNA and acpcPNA-RNA duplexes more closely resembled the typical structures of B-DNA and A-RNA than the corresponding duplexes of aegPNA. The calculated binding free energies are in good agreement with the experimental results that the acpcPNA-DNA duplex is more stable than the acpcPNA-RNA duplex regardless of the base sequences. The results provide further insights in the relationship between structure and stability of this unique PNA system.

1. Introduction

Since 2001, various pyrrolidinyl peptide nucleic acids (PNAs) have been synthesized by Vilaivan and coworkers [1–7]. Pyrrolidine rings and other cyclic spacers have been incorporated into novel PNA backbones to reduce the backbone flexibility of the original *N*-2-aminoethylglycine PNA (aegPNA) [8]. acpcPNA, a representative pyrrolidinyl PNA, has a unique structure consisting of nucleobase-modified D-proline with a (2′*R*,4′*R*)-configuration and (1*S*,2*S*)-2-aminocyclopentane carboxylic acid [4] (Fig. 1). The (2′*R*,4′*S*)-proline epimer of acpcPNA was later designed [5,6]. Both acpcPNA and its epimer (*epi*-acpcPNA) can form highly stable hybrid duplexes with DNA and RNA. The thermal stability of the duplexes was determined by measuring their melting temperatures (T_m), and acpcPNA-DNA duplexes were found to have considerably higher melting temperatures and to be more sensitive to a mismatched Watson-Crick base pairing than aegPNA-DNA duplexes [7–11]. In contrast to aegPNA, the binding affinity of acpcPNA for DNA is significantly higher than that for RNA [6]. Due to the advantages in efficiency and sequence specificity of pyrrolidinyl PNA for binding DNA and RNA, it could potentially be developed and used in a wide variety of applications, including sensor probes, self-reporting fluorescence probes and DNA sensing based on the differential electrostatic properties of PNA and DNA [7].

The availability of three-dimensional structures of aegPNA, including PNA-PNA, PNA-DNA and PNA-RNA duplexes as well as a PNA-DNA-PNA triplex, allows researchers to study their properties on a molecular level (for X-ray, NMR and theoretical data, see, for example, Refs. [12–18]). The experimental structure of pyrrolidinyl PNA, in contrast, is unavailable, and only limited structures derived from modeling of acpcPNA and *epi*-acpcPNA are available [19–22]. Therefore, experiments that provide insight into the structural properties and molecular behavior of this pyrrolidinyl PNA are needed. Previously, we reported the structural and energetic properties of the self-complementary and DNA hybrids of pyrrolidinyl PNA as investigated by molecular dynamics (MD) simulations [20–22]. The MD structures of PNA self-hybrids exhibit B- and P-type conformations depending on the relative orientations of the two strands [22]. On the other hand, the conformations of both acpcPNA-DNA and *epi*-acpcPNA-DNA duplexes are close to the B-

* Corresponding author.
E-mail address: skhatcha@kku.ac.th (K. Siriwong).

https://doi.org/10.1016/j.jmgm.2018.06.008
1093-3263/

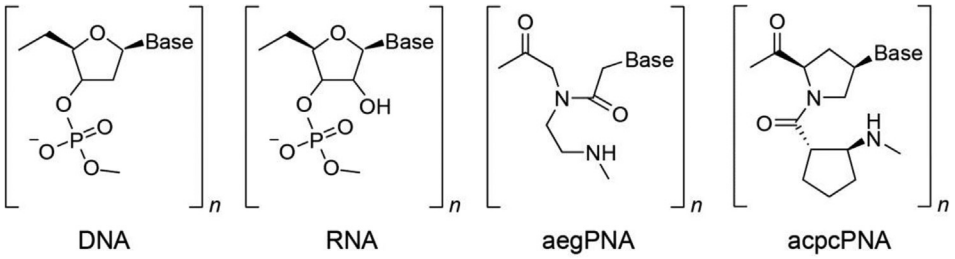

Fig. 1. Monomeric units of DNA, RNA, and PNA with *N*-2-aminoethylglycine (aegPNA) and D-prolyl-(1*S*,2*S*)-2-aminocyclopentane carboxylic acid (acpcPNA) backbones.

form [21], indicating that the pyrrolidinyl PNA strand can mimic the conformation of its complementary DNA strand. The calculated thermodynamic stability is in the order of acpcPNA-DNA > *epi*-acpcPNA-DNA > DNA-DNA, which is in good agreement with the T_m data [20,21], where gas phase enthalpy is the main contributor to duplex stability.

To the best of our knowledge, theoretical studies of the pyrrolidinyl PNA-RNA system have yet to be performed. In this work, MD simulations of acpcPNA-RNA and acpcPNA-DNA duplexes with different base sequences were carried out in an aqueous solution to determine the structural and binding properties of acpcPNA bound to RNA and to compare the binding of acpcPNA to RNA and DNA.

2. Materials and methods

2.1. Molecular dynamics simulations

Four MD simulations of PNA-DNA and PNA-RNA duplexes with two different base sequences were performed using the sander module of the AmberTools 17 program [23]. The base sequences of the PNA strand were *N*-AGTGATCTAC-*C* (**P1**) and *N*-GCGACGTAGC-*C* (**P2**), and these PNA strands bound to their complementary DNA and RNA strands in an antiparallel fashion; i.e., the N-terminus of the PNA strand was oriented toward the 3′-end of the complementary DNA/RNA strands. These base sequences were chosen due to their diverse % (G + C) content (40 and 70% for the **P1** and **P2** systems, respectively) and the availability of experimental thermal stability data [6], which allowed us to investigate the effect of % (G + C) content on duplex stability and compare with simulated results. The starting PNA-DNA and PNA-RNA duplex structures were constructed as follows. First, the DNA-DNA duplex was generated using the NAB module of the AmberTools 17 program based on a canonical B-DNA (rise parameter of 3.4 Å, twist angle of 36° and approximately zero for other helical parameters). The RNA-RNA duplex was generated with the same procedure except for a 32° twist angle, which is based on the Arnott A-RNA conformation [24]. Next, the DNA/RNA nucleotides in the first strand of these duplexes were replaced by the corresponding PNA monomeric units prepared previously [20,21] with fixed nucleobase coordinates. Finally, the PNA backbone units were linked to each other and then re-optimized with the molecular mechanics method.

MD simulations were carried out in an aqueous solution with the ff99bsc0+OL15 force field [25,26] for PNA-DNA and ff99bsc0+chiOL3 [27] for PNA-RNA systems. All force field parameters (for bonds, angles, torsions, etc.) for the PNA unit were also taken from the ff99 force field, only their atomic point charges were generated using the restrained electrostatic potential (RESP) method as described previously [20]. The negative charge of the DNA or RNA strand was neutralized by 9 sodium ions, and the duplex was embedded into a rectangular box of the TIP3P water model [28] extended by 10 Å in each direction from the duplex. The simulations were performed in the isothermal-isobaric ensemble ($T = 298.15$ K and $P = 1$ atm) under standard conditions (time step of 2 fs, periodic boundary conditions, 9 Å cut off for nonbonded interactions; the particle mesh Ewald method was employed for long-range interactions [29], and the SHAKE algorithm was employed for constraining all hydrogen atom movement [30]). The simulation protocol for the heating and equilibration steps followed the standard procedures described elsewhere [20]. After achieving equilibration, the production runs were carried out for 60 ns with unrestrained MD simulations, and the MD structures were collected every 2 ps (totaling 30,000 structures for each MD trajectory). The last 50 ns MD trajectories were employed for further analysis.

2.2. Binding free energy calculations

The most widely used computational method to evaluate the binding affinity between PNA and DNA/RNA strands is the molecular mechanics/generalized Born surface area (MM-GBSA) approach [19,20,31]. The binding free energy ($\Delta G_{binding}$) was calculated as

$$\Delta G_{binding} = \Delta H_{gas} + \Delta G_{solv} - T\Delta S \quad (1)$$

where Δ denotes the difference in energy between the duplex and the individual PNA and DNA/RNA strands; i.e., $\Delta x = x_{duplex} - (x_{PNA} + x_{DNA\ or\ RNA})$, where x is the energy term. H_{gas} is the gas phase energy, which is the summation of the internal energy (bond, angle and dihedral, E_{int}), and electrostatic (E_{elec}) and van der Waals (E_{vdW}) interactions obtained from MM calculations. The solvation free energy, G_{solv}, is the summation of the polar (G_{polar}) and nonpolar ($G_{nonpolar}$) terms solved by using the GB model and the solvent accessible surface area (SASA) method, respectively. The last term is the product of the absolute temperature T (298.15 K) and entropy S estimated by normal-mode analysis.

3. Results and discussion

3.1. Structural properties of simulated duplexes

The conformational stability of the simulated duplexes was investigated in terms of the root-mean-square deviation (RMSD), which reflects differences in the geometry of the simulated structures from their starting structures. All duplex atoms were considered in the RMSD calculations. Overall, the MD trajectories of all duplexes were stable in terms of geometry during a 60 ns simulation time; see Fig. 2a. Interestingly, the RMSDs of PNA-RNA duplexes were relatively small (less than 2 Å) whereas those of PNA-DNA duplexes were notably larger (approximately 3 Å), indicating a slightly increased divergence of PNA-DNA duplexes from their starting structures compared to PNA-RNA systems. The RMSDs were also investigated with respect to MD-averaged