THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

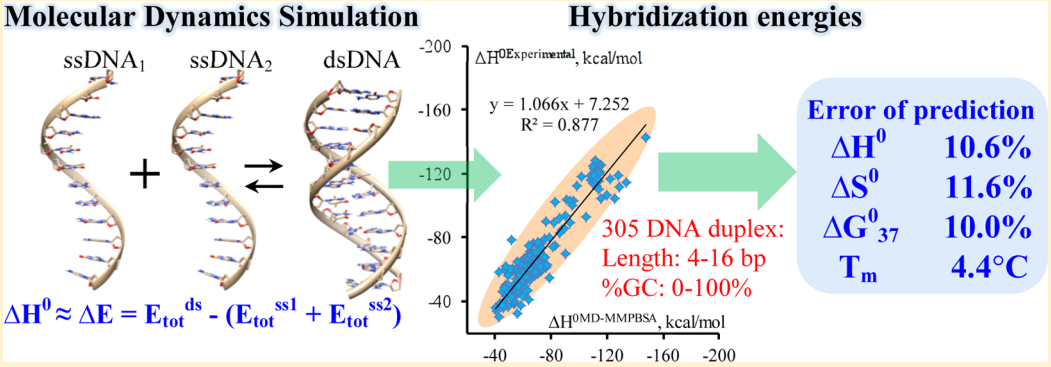
Evaluation of the Gibbs Free Energy Changes and Melting Temperatures of DNA/DNA Duplexes Using Hybridization Enthalpy Calculated by Molecular Dynamics Simulation

Alexander A. Lomzov,*[,†,‡] Yury N. Vorobjev,[†,‡] and Dmitrii V. Pyshnyi*[,†,‡]

[†]Institute of Chemical Biology and Fundamental Medicine, Novosibirsk 630090, Russia

[‡]Novosibirsk State University, Novosibirsk 630090, Russia

Supporting Information



ABSTRACT: A molecular dynamics simulation approach was applied for the prediction of the thermal stability of oligonucleotide duplexes. It was shown that the enthalpy of the DNA/DNA complex formation could be calculated using this approach. We have studied the influence of various simulation parameters on the secondary structure and the hybridization enthalpy value of Dickerson−Drew dodecamer. The optimal simulation parameters for the most reliable prediction of the enthalpy values were determined. The thermodynamic parameters (enthalpy and entropy changes) of a duplex formation were obtained experimentally for 305 oligonucleotides of various lengths and GC-content. The resulting database was studied with molecular dynamics (MD) simulation using the optimized simulation parameters. Gibbs free energy changes and the melting temperatures were evaluated using the experimental correlation between enthalpy and entropy changes of the duplex formation and the enthalpy values calculated by the MD simulation. The average errors in the predictions of enthalpy, the Gibbs free energy change, and the melting temperature of oligonucleotide complexes were 11%, 10%, and 4.4 °C, respectively. We have shown that the molecular dynamics simulation gives a possibility to calculate the thermal stability of native DNA/DNA complexes a priori with an unexpectedly high accuracy.

INTRODUCTION

Oligonucleotides (short synthetic fragments of nucleic acids) are widely used in various fields of fundamental and applied research. To make the design and applications of oligonucleotide derivatives and analogues more effective, their properties should be thoroughly investigated. There are two essential physicochemical properties of oligomeric nucleic acids (NAs) which determine their functions: secondary structure and the complex formation ability. As a rule, the thermal stability of complexes is predicted with the use of the nearest neighbor approach. The enthalpy ($\Delta H°$), entropy ($\Delta S°$), or Gibbs free energy ($\Delta G°(T)$) of the duplex formation are calculated as the sum of the individual nearest neighbors impacts (thermodynamic increments) and corrections that characterize the formation of structural elements of the double helix (dinucleotide base pairs and terminal base pairs) and the symmetry of complexes (e.g., refs 1−3). An estimation of these increments requires the synthesis of a series of model oligonucleotides followed by detailed analysis of the spatial organization of their complexes and their hybridization properties. It should be noted that the synthesis of a set of new oligonucleotide derivatives and analogues is a laborious task, and it is not always possible to predict their physicochemical characteristics reliably. The possibility to perform a prognostic calculation of NA secondary structure and thermal stability by computer simulation would greatly facilitate the design of new molecular constructs with predetermined characteristics.

Due to recent developments in computer hardware and the progress in computer simulations, the molecular dynamics

Received: October 2, 2015
Revised: November 16, 2015
Published: November 16, 2015

DOI: 10.1021/acs.jpcb.5b09645
J. Phys. Chem. B 2015, 119, 15221−15234

(MD) method is increasingly being applied for *in silico* simulation of the properties of biopolymers. This method is often used for the characterization of physicochemical properties of double-stranded nucleic acids, e.g., the detailed characterization of the structure of the polyA/polyT regions,[4] the examination of the B → A and B → Z transitions of the double helix in DNA,[5,6] the determination of conformational flexibility of structural elements of nucleic acids,[7,8] studies of the nucleotide sequence-dependent solvation process,[9] and the interactions of nucleic acids with cations[10−12] or ligands.[13] Computer simulation methods are also used to characterize the properties of artificial oligonucleotide derivatives. For example, the structure and dynamics of modified oligomer complexes containing locked nucleic acids (LNAs),[14,15] peptide nucleic acids (PNAs),[16] nucleobase modifications,[17] and terminal DNA duplex modifications[18] were characterized. Simulation time was achieved the microsecond range.[19] Nevertheless, realistic MD simulations still face a lot of challenges.[20] Among them, we should mention a careful choice of force fields and force field parameters as being the best or most suitable (for example, see ref 21), etc.

A majority of computer simulations are focused on the characterization of ternary structures. On the contrary, this methodology can be used for the predictive calculations of the efficiency of biopolymers' interactions. In this case, it is necessary to determine the energy differences between various states of biomolecules, and not the absolute energy values of the states, which may vary by several orders of magnitude. First attempts of such calculations have already been undertaken.[22−26] However, to the best of our knowledge, the possibility to calculate thermodynamic parameters of the complex formation (enthalpy, entropy, and Gibbs free energy changes) has not been studied in detail in the case of nucleic acids.

The goal of this work was to examine the applicability of the molecular dynamics method for the calculation of thermal stability of DNA duplexes. To achieve this purpose, the study included two stages: (1) the calculation of hybridization enthalpy for native oligomers, and (2) the estimation of Gibbs free energy changes and melting temperatures of DNA duplexes.

MATERIALS AND METHODS

Molecular Dynamics Simulation. MD simulations were performed using the Amber 12 software package (AmberTools 12 and Amber 12), with the use of parallel calculations on central (CPU) and graphical (GPU) processors. The simulations were performed on a personal computer with the following configuration: CPU Intel i7 2600K, 4 GB memory, and GTX 580 video card. In addition, the resources of the Siberian Super Computer Center (Novosibirsk, Russia) were employed for the simulation of oligonucleotides included in the database.

The study was carried out using the ff99bsc0 force field[27] which is one of the best for the simulation of B-DNA.[28] SHAKE algorithm for hydrogen involving bonds was applied. Long range electrostatics was calculated using particle mesh Ewald (PME), with a 1 Å grid.

Molecular dynamic simulations in implicit solvent were performed at constant temperature using Anderson (strong collision) temperature regulation scheme. A modified GB model with 0.1 M concentration (unless otherwise specified) of 1-1 mobile counterions in solution was used.[29,30] A time step of 1 fs was used. Coordinates of each atom of the system (snapshots) were saved every 1 ps.

A productive molecular dynamics trajectory in explicit solvent was generated in the NPT ensemble. The pressure coupling algorithm of the "weak-coupling" variety with isotropic position scaling analogous to temperature coupling was used.[31] The pressure of the system was maintained at 1 bar. The usage of the SHAKE algorithm allows setting a 2 fs time step. Anderson (strong collision) temperature regulation scheme was used. Nonbonded cutoff of 10 Å was applied. Coordinates of each atom of the system (snapshots) were saved every 1 ps.

Weak harmonic potential for heavy atoms in MD simulation (0.001 kcal/mol/Å^2) was used to prevent compactization of single-stranded oligonucleotides.

Simulation in the Implicit Solvent (Generalized Born Model). The simulation included six stages: (1) generation of the pdb file containing the structure of DNA duplex in the B form (NAB, AmberTools 12; the single strands were prepared from double-stranded DNA), (2) conversion of the pdb file to the Amber7 format (tleap, AmberTools 12), (3) minimization of the energy of DNA (PMEMD.MPI, Amber 12), (4) heating to the desired temperature for 2.5 ns (PMEMD.CUDA, Amber 12), (5) equilibration of the system for 1 ns at constant temperature (PMEMD.CUDA, Amber 12), (6) MD simulation for 10 ns (PMEMD.CUDA, Amber 12).

Simulation in the Explicit Solvent Included 13 Stages. The first 5 stages of simulation in the explicit solvent were the same as those in the implicit water shell. The remaining stages were the following: (6) creation of the pdb file of the equilibrated structure (ambpdb, AmberTools 12), (7) addition of the water shell (TIP3P water, cuboid periodic box, 12 Å) and sodium ions to neutralize the net charge of DNA and conversion to the Amber7 format (tleap, AmberTools 12), (8) minimization of all systems with fixed DNA (10 000 steps, constant restraint force of 500 cal/mol/Å^2) (PMEMD.MPI, Amber 12), (9) heating of all the system with fixed DNA for 2.5 ns with time step of 0.0005 ps (constant restraint force of 500 cal/mol/Å^2) (PMEMD.CUDA, Amber 12), (10) equilibration of the system density at the constant volume for 500 ps (PMEMD.CUDA, Amber 12), (11) equilibration of the system density at the constant pressure (1 bar) for 500 ps (PMEMD.CUDA, Amber 12), (12) MD productive trajectory simulation, 10 ns (PMEMD.CUDA, Amber 12) at NTP ensemble, (13) energy calculation (MMPBSA, AmberTools 12).

Calculation of Hybridization Enthalpy. Hybridization enthalpy (ΔH°) was calculated as a difference of the oligonucleotides' total energy in the double-stranded (ds) state (E_{tot}^{ds}) and in the single-stranded (ss) state (E_{tot}^{ss1} and E_{tot}^{ss2}):

$$\Delta H^\circ \approx E_{tot}^{ds} - (E_{tot}^{ss1} + E_{tot}^{ss2}) \quad (1)$$

The value of hybridization enthalpy was close to the total internal change because the simulation was performed at constant pressure, and the change of the volume during DNA denaturation was negligible.

The total energy (E_{tot}^{i}, i = ds, ss1, ss2) of each state was taken as the average value along the MD trajectory after an equilibration. The error of the total energy calculation was calculated as the root-mean-square deviation (RMSD) of total energy along the trajectory divided by the square root of the

DOI: 10.1021/acs.jpcb.5b09645
J. Phys. Chem. B 2015, 119, 15221−15234

Table 1. Comparison of the MD Simulation Results Using Various Parameters for the Dickerson−Drew Dodecamer

thermostat type and conditions	heating time, ns	T, K	[Na^+], M	$RMSD_{ds1}$, Å	$RMSD_{ds2}$, Å	$\Delta H°$, kcal/mol	E_{tot}^{ds}, kcal/mol	E_{tot}^{ss}, kcal/mol
				Stage 1				
Andersen[a]	2.5	300	0.1		2.03	−114.3 ± 7.92	−4097.1 ± 3.6	−1991.9 ± 2.17
Andersen[b]	2.5	300	0.1		2.04	−113.4 ± 7.92	−4097.0 ± 2.9	−1991.9 ± 2.17
				Stage 2				
Andersen	0.1	300	0.1	0.05	2.05	−122.1 ± 5.01	−4097.3 ± 2.9	−1987.6 ± 2.1
Andersen	0.25	300	0.1	0.07	2.05	−105.7 ± 4.97	−4097.8 ± 2.9	−1996.1 ± 2.1
Andersen	0.5	300	0.1	0.07	2.06	−113.4 ± 5.01	−4097.5 ± 2.9	−1992.1 ± 2.1
Andersen	1	300	0.1	0.07	2.05	−115.8 ± 4.99	−4097.8 ± 2.9	−1991 ± 2.1
Andersen	5	300	0.1	0.07	2.03	−107.4 ± 4.99	−4097.9 ± 2.9	−1995.3 ± 2.1
Andersen	10	300	0.1	0.06	2.04	−117.4 ± 4.98	−4098.2 ± 2.9	−1990.4 ± 2.1
Andersen	100	300	0.1	0.07	2.01	−100.5 ± 4.96	−4096.6 ± 2.9	−1998.1 ± 2.1
				Stage 3				
Andersen	2.5	1	0.1		0.11	−110.2 ± 0	−5466 ± 0.0	−2677.9 ± 0.0
Andersen	2.5	10	0.1	0.94	0.36	−109.4 ± 0.3	−5425.4 ± 0.1	−2658 ± 0.1
Andersen	2.5	50	0.1	0.90	0.77	−109.2 ± 1.1	−5244 ± 0.5	−2567.4 ± 0.3
Andersen	2.5	100	0.1	0.8	1.08	−111.8 ± 2.3	−5016.7 ± 0.9	−2452.5 ± 0.7
Andersen	2.5	150	0.1	0.65	1.32	−115.4 ± 3.5	−4788.9 ± 1.4	−2336.7 ± 1.0
Andersen	2.5	200	0.1	0.51	1.57	−115.3 ± 5.1	−4557.9 ± 2.4	−2221.3 ± 1.4
Andersen	2.5	250	0.1	0.26	1.79	−117.0 ± 6.4	−4327.7 ± 3.0	−2105.4 ± 1.7
Andersen	2.5	300	0.1		2.06	−114.3 ± 7.9	−4098.1 ± 3.6	−1991.9 ± 2.2
Andersen	2.5	350	0.1	0.36	2.59	−114.6 ± 9.1	−3861.4 ± 4.3	−1873.4 ± 2.4
Andersen	2.5	400	0.1	3.92	6.00	−109.0 ± 10.5	−3594 ± 4.9	−1742.5 ± 2.8
				Stage 5				
Andersen	2.5	300	0.01	0.31	2.03	−100.3 ± 7.8	−4074.5 ± 3.6	−1987.1 ± 2.1
Andersen	2.5	300	0.5	0.28	2.03	−120.1 ± 7.2	−4113.1 ± 2.9	−1996.5 ± 2.1
Andersen	2.5	300	1	0.28	2.05	−124.1 ± 7.0	−4117.9 ± 2.9	−1996.9 ± 2.1
				Stage 6				
Langevin, 1 ps^{-1}	2.5	300	0.1	0.07	2.07	−91.5 ± 7.1	−4077 ± 2.9	−1992.8 ± 2.1
Langevin, 2 ps^{-1}	2.5	300	0.1	0.12	2.09	−102.6 ± 7.2	−4092 ± 2.9	−1994.7 ± 2.2
Langevin, 5 ps^{-1}	2.5	300	0.1	0.15	2.14	−104.4 ± 7.1	−4091.2 ± 2.9	−1993.4 ± 2.1
Langevin, 10 ps^{-1}	2.5	300	0.1	0.16	2.16	−115.8 ± 7.1	−4091.5 ± 2.9	−1987.8 ± 2.1
Langevin, 25 ps^{-1}	2.5	300	0.1	0.44	2.07	−120.6 ± 7.1	−4090.7 ± 2.9	−1985.0 ± 2.1
Langevin, 50 ps^{-1}	2.5	300	0.1	0.47	2.02	−129.1 ± 7.0	−4106.7 ± 2.9	−1988.8 ± 2.1
				Stage 7				
water, GB, 6 Å[c]	2.5	300	N			−120.8 ± 3.2	−2937.8 ± 1.4	−1408.5 ± 0.1
water, PB, 6 Å	2.5	300	N	1.16	2.26	−96.4 ± 3.3	−3030.9 ± 1.4	−1467.2 ± 0.1
water, GB, 8 Å	2.5	300	N			−119.6 ± 3.1	−2937.6 ± 1.4	−1409.0 ± 0.9
water, PB, 8 Å	2.5	300	N	1.34	2.21	−94.5 ± 3.3	−3031.8 ± 1.4	−1468.7 ± 0.9
water, GB, 10 Å	2.5	300	N			−120.0 ± 3.4	−2936.6 ± 1.4	−1408.3 ± 1.0
water, PB, 10 Å	2.5	300	N	0.98	2.22	−97.2 ± 3.5	−3031.7 ± 1.4	−1467.3 ± 1.0
water, GB, 12 Å	2.5	300	N			−117.0 ± 3.7	−2925.8 ± 1.4	−1404.4 ± 1.0
water, PB, 12 Å	2.5	300	N	2.44	2.33	−93.8 ± 3.7	−3021.7 ± 1.4	−1463.9 ± 1.0
water, GB, 14 Å	2.5	300	N			−121.4 ± 3.4	−2942.1 ± 1.4	−1410.4 ± 1.0
water, PB, 14 Å	2.5	300	N	1.41	2.33	−96.9 ± 3.5	−3035.5 ± 1.4	−1469.3 ± 1.0
water, GB, 16 Å	2.5	300	N			−118.6 ± 3.4	−2926 ± 1.5	−1403.7 ± 1.0
water, PB, 16 Å	2.5	300	N	1.04	2.37	−93.6 ± 3.4	−3021.6 ± 1.5	−1464.0 ± 1.0
		16		3.68[d]				
		290		3.47[e]				
				3.41[f]				

[a]Results were obtained using CPU on one core. [b]Results were obtained using CPU in parallel mode on 8 cores. All other results were obtained using GPU. N designates that charge was neutralized by adding sodium ions. [c]Simulations were performed in Langevin thermostat (2 ps^{-1}) in the explicit water shell with specified size in Å. [d]PDB ID: 2BNA, obtained by X-ray diffraction.[38] [e]PDB ID: 7BNA obtained by X-ray diffraction.[39] [f]PDB ID: 2DAU, obtained by NMR method.[41] PB: using MMPBSA to calculate energy. GB: using MMGBSA to calculate energy. All other simulations were performed in implicit solvent shell.

number of independent snapshots minus 1. For details see Supporting Information.

In the case of the direct enthalpy calculations in the implicit water shell, the total internal energy of single- and double-stranded states was obtained in the course of an independent simulation of each state.

In the case of explicit solvent, the hybridization enthalpy values were obtained by analyzing the MD trajectory only of the dsDNA using Molecular Mechanics Generalized Born (or Poisson−Boltzmann) Surface Area (MMGB(PB)SA) calculations. In other words, single-stranded oligonucleotides were "withdrawn" from the trajectory of double-stranded state.

DOI: 10.1021/acs.jpcb.5b09645
J. Phys. Chem. B 2015, 119, 15221−15234

Thermal Denaturation Studies. The melting of oligonucleotide complexes was carried out in a thermoregulated cell equipped with a special device based on the UV detector of Milichrom liquid chromatograph (Econova, Russia), as described in ref 32. Each melting curve comprised at least 600 absorbance values with a frequency of 10 points/°C and was recorded with the heating rate of 0.7−1 °C/min. Heating curves coincided with cooling curves, which indicated that the formation of the complexes was under thermodynamic equilibrium conditions. The recording of the melting curves in a multiwavelengths mode was carried out by an automatic switching of the monochromator between four wavelengths in the range 230−300 nm with 10 nm steps. Integration time for the absorbance value at each wavelength did not exceed 1.2 s. Complexes were formed by stoichiometric mixing of the oligonucleotide strands in the range of total concentration of 1−10 μM. Melting curves were fitted according to the two-state model as described in ref 33.

Thermal denaturation studies were carried out in a buffer containing 1 M NaCl, 10 mM sodium phosphate, and 0.1 mM sodium EDTA (pH 7.2).

Database of Thermodynamic Parameters. The database of experimental values concerning the thermal stability of native fully complementary duplexes of oligodeoxyribonucleotides (enthalpy, entropy, and Gibbs free energy change) included thermodynamic parameters of 305 complexes with the length of 4−16 bp (average 9.2) and GC content of 0−100% (average 56%). The data were obtained by thermal denaturation method (28 complexes) or taken from the literature (277 complexes) (Supporting Information, Table S1).[34−36]

RESULTS

I. Hybridization Enthalpy Calculation. *Search for Optimal Simulation Parameters.* Molecular dynamic simulation of biopolymer properties implies a significant number of parameters to be set, including the force field, the heating protocol, the simulation temperature, the type of water shell (explicit or implicit), the concentration of ions in solution, the simulation time, etc.[37]

At the first stage of this study, we determined the optimal simulation parameters, and then performed a simulation of a representative set of DNA duplexes to obtain the enthalpy changes of the corresponding complex formation. To evaluate the optimal simulation parameters for the predictive calculation of the thermal stability of DNA complexes, the Dickerson−Drew dodecamer (DDD) (d(CGCGAATTCGCG)$_2$) was used. Its secondary structure[38−43] and thermal stability are well-studied, and thermodynamic parameters of the bimolecular complex formation (the enthalpy and entropy change) were determined previously.[44]

The search for the optimal simulation parameters was divided into 8 stages which included the variation of the following parameters: (1) usage of CPU and GPU for simulation, (2) heating time, (3) simulation temperature, (4) simulation time after equilibration, (5) ionic strength in the implicit water shell, (6) temperature control method (in the case of the implicit water shell), (7) the water shell radius.

The main reliability criteria for the MD simulation of oligonucleotides and their complexes are the preservation of their double-stranded state and the spatial structure close to the B-form of DNA. The third criterion is the close agreement of experimental hybridization enthalpy change with that determined using MD simulation.

1. Use of GPU for MD Simulation. At the first step, we estimated a possibility of using graphics processors for MD simulation. The comparison of simulation results in the implicit solvent using central and graphics processors is presented in Table 1, stage 1. The simulation was performed in the Anderson thermostat at 300 K and 0.1 M concentration of monovalent cations. Average structures of the dodecamer complexes calculated with the use of either CPU in a one core mode and in a parallel mode on 8 cores or GPU coincided with each other. Nearly the same values were obtained for $RMSD_{1ds}$, which is the root-mean-square deviation between the structure averaged for 10 ns trajectory at 300 K in Andersen thermostat and the structure averaged along the trajectory at the other choices of simulation setup or obtained experimentally (Table 1). Values of $RMSD_{2ds}$, which were determined as standard deviations of heavy atoms from the equilibrium position along MD trajectories, were almost equal in both cases (Table 1). In addition, total energy values for the oligonucleotide (E_{tot}^{ss}) and the complex (E_{tot}^{ds}) (Table 1) calculated with both types of processors were also similar. The internal energy of complex formation evaluated as the difference between total energies of double- and single-stranded states (Figure 1, eq 1) was close to

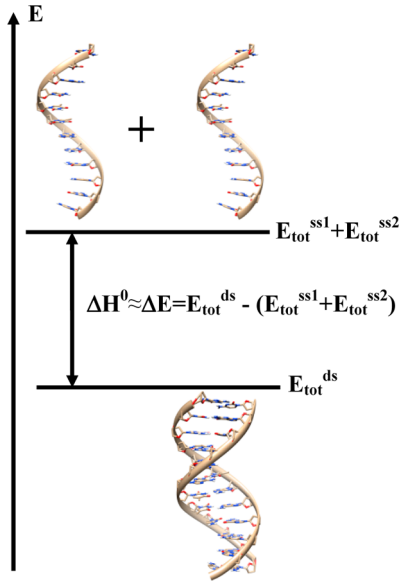


Figure 1. Scheme for the calculation of the enthalpy for DNA duplex formation.

the corresponding ΔH° value and did not depend significantly on simulation method. Moreover, enthalpy change values of the complex formation obtained by the MD method (the average value for all three methods was 113.8 kcal/mol) were close to experimental values (−116 kcal/mol)[44] and those calculated using the nearest neighbor model (NN) (−95.5 ± 9.5 kcal/mol).[3]

It should be noted that the use of parallel calculations allows us to significantly accelerate the calculations by MD method. The development of computer hardware gives a possibility to perform the simulation at the desktop computer, and the employment of graphics processors increases the speed of calculations as compared to that from the central processor alone. For example, the efficiency of the Dickerson−Drew dodecamer simulation in the implicit water shell using a single core was 1.4 ns/day versus 6.32 ns/day for 8 cores. A further

DOI: 10.1021/acs.jpcb.5b09645
J. Phys. Chem. B 2015, 119, 15221−15234

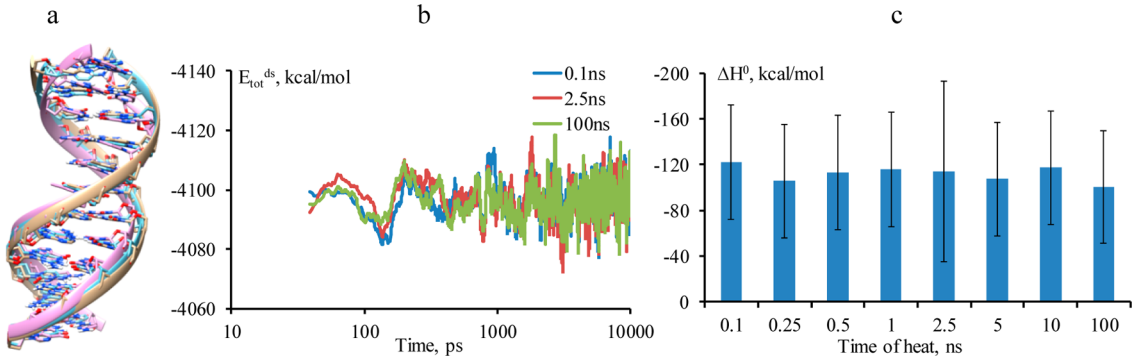


Figure 2. Analysis of 10 ns MD trajectories after equilibration. (a) The comparison of structures in double-stranded state averaged along MD trajectory in explicit (12 Å) (blue) and implicit (brown) water shell at the heating time of 2.5 ns with the structure obtained by the X-ray crystallography method (pink).[39] (b) Evolution of the Dickerson−Drew dodecamer duplex total energy after the equilibration at various heating times of the system in the implicit water shell. (c) The comparison of the enthalpy values of complex formation at various heating times in the implicit water shell.

increase up to 222 ns/day was achieved with the use of GPU GTX 580. In the case of the explicit water shell, the use of GPU GTX 580 allows calculation of the MD trajectory with the length of 39 ns/day. Thus, even a personal computer provides a possibility to simulate and calculate the energy of the formation of two DNA complexes per day.

To summarize, the results of the MD simulation obtained with the use of central and graphics processors are equivalent. This allows one to use graphics processors to significantly reduce the computation time.

2. Heating Time. At the second stage, we studied the influence of the heating time of the MD system on simulation results in implicit solvent. The studied system was heated from 1 to 300 K linearly over 0.1−100 ns. In all cases, structures of DNA complexes averaged along the equilibrium region of the MD trajectory were close to each other (see the $RMSD_{1ds}$ values in Table 1, stage 2). It was observed that the double DNA helix was preserved, its topology was close to the B-form of DNA double helix, and the structures obtained by the MD calculations were nearly the same (Figure 2a). At the same time, all the structures obtained by the computer simulation method were slightly different from the structures obtained by the X-ray crystallography method: the $RMSD_{1ds}$ value was on average ∼3.5 Å. The close values of the $RMSD_{1ds}$ are reported in the literature. For example, the analysis of the 60 ns MD simulation in the AMBER force field parm94 shows similar RMSD values of the calculated structures and a number of structures obtained using X-ray (2.8 ± 0.7 Å) and NMR (2.6 ± 0.5 Å) techniques.[42] In the resent studies, the comparison of MD simulation in AMBER force field (parambsc00) with X-ray data (PDB ID: 7BNA) shows the values of RMSD 4.06 Å.[43]

Figure 2b shows that the equilibrium parts of molecular dynamic trajectories are close to each other in the range 0.025−10 ns, independent of the heating time of the system (0.1−100 ns). Moreover, energies of complexes in the single-stranded state and their difference (corresponding to the hybridization enthalpy change) (Figure 2c) coincide with the accuracy, which is much better than the calculation error. We have chosen 2.5 ns as a typical heating time from 1 K up to 300 K.

3. Simulation Temperature. Studies of the effect of simulation temperature on the duplex structure and hybridization energy have shown that the temperature increase leads to the increase in the flexibility of structure, as indicated by $RMSD_2$ values calculated along the MD trajectory (Table 1). With the rise of the temperature, the $RMSD_{2ds}$ value increases linearly within the temperature range 1−350 K (Table 1, stage 3; Supporting Information, Figure S1). The flexibility of the DNA duplex structure along the trajectory significantly increases at 400 K ($RMSD_{2ds}$ = 6.0 Å). The averaged conformation of the duplex at 400 K also differs from that at 300 K ($RMSD_{1ds}$ = 3.92 Å). These effects are caused by unwinding of terminal pairs and/or base flipping from the duplex which were observed experimentally[34−48] and studied using MD simulation.[49]

Average structures of ssDNA do not depend on the temperature in the 0−350 K range (values $RMSD_1$ = 1.46 ± 0.3 Å) due to weak restrained potential (0.001 kcal/mol/A^2) applied for heavy atoms. At the same time, $RMSD_2$ linearly increases in all studied ranges from 0.11 to 2.59 Å (Supporting Information, Figure S1).

Thus, the structures of oligonucleotides and their complexes averaged along MD trajectory at different temperatures from 1 to 350 K are close to each other. Consequently, the enthalpy change values of the complex formation should also be similar. Internal DNA energies in single-stranded and double-stranded states vary by a factor of 1.5. For instance, the internal energy of dsDNA increases from −5466.1 to −3594 kcal/mol with the temperature rise from 1 to 400 K (Table 1). Nevertheless, the corresponding values of the duplex formation enthalpy change coincide within the given temperature range (ΔH° vary from −110.2 to −117 kcal/mol), and ΔH° variations along MD trajectory increase with increasing temperature (from 0.7 to 105 kcal/mol, Table 1). A temperature of 300 K was chosen for simulation since it is close to the physiological value.

4. Simulation Time after Equlibration. The next important parameter, which can influence the internal energy values, is the time of simulation after an equilibration of the molecular dynamic system. The number of microstates of the system may be not sufficient at short simulation times, thus compromising the reliability of the mean values of the system's internal energy. As a result, the obtained internal energy values would not correspond to the values of the equilibrated thermodynamic system. Therefore, the former values would not be suitable for the calculation, and cannot be compared with experimental values of double-stranded to single-stranded state transition energy. On the other hand, it is not reasonable to carry out the simulation for too long a period of time, since the average energy value became constant after a certain simulation time.

DOI: 10.1021/acs.jpcb.5b09645
J. Phys. Chem. B 2015, 119, 15221−15234

The study of the influence of the simulation time on energy values of complexes and the complex formation revealed that single-stranded oligonucleotides fold into a compact coil structure at the times over the 10 ns, if the harmonic restraint potential was not applied (Figure S2).

However, the experimental data obtained by a number of methods such as small-angle X-ray scattering (SAXS),[50,51] fluorescence resonance energy transfer (FRET),[51−53] gel electrophoresis,[54] atomic force microscopy,[55] etc. gives the values 1−4 nm (0.5−1 turn of the helix). For example, the values of ∼1 nm are reported for poly(dT) in 0.1 M sodium salt at room temperatures.[50] For oligo(dA) sequences with more efficient base stacking, the persistence length is higher (∼3−6 nm).[50] Thus, persistence length of the oligonucleotides with heterogenic nucleotide content (various stacking in the single-stranded chain) under the above-mentioned conditions is close to the 10 nucleotides. Besides, the backbone of ssDNA usually has the conformation resembling the B-form of dsDNA, as confirmed by circular dichroism spectra at various temperatures for oligonucleotides (for example, circular dichroism data in ref 33 and reference therein), polymers, and dinucleotide.[56] As a rule, stacking interactions within oligomer sequences are fully disrupted at temperatures above 50−70 °C. So, the single-stranded Dickerson−Drew 12-mer oligonucleotide should not be a collapsed structure. In addition, despite the potential ability of DDD to form an intermolecular hairpin, 1 μs time was not enough to form an ordered single-stranded structure.

To avoid the compaction during the single-stranded structure simulation, we imposed a weak harmonic potential on heavy atoms of the main chain of the oligodeoxyribonucleotide molecules. To minimize the effect of restraints on hybridization energy changes, we have used a weak harmonic potential (0.001 kcal/mol/Å^2) for heavy atoms of nucleic acids in single- and double-stranded states for the simulation in implicit water shell. The total restraint energy penalty averaged along the MD trajectory was ∼8 kcal/mol which is negligible as compared to total internal energy (∼2000 kcal/mol).

This penalty is the probability of nonideality of the force field which was developed for dsDNA.

The structure of the double helix did not change throughout the simulation time (1000 ns). Structures averaged for various periods of times from 0.5 to 1000 ns, and the values of standard deviations of atoms from the structure with minimized total energy in the implicit water shell, are shown in Supporting Information Figure S3. The analysis of the hybridization energy effect on the simulation time demonstrates that, after 1 ns equilibration of the system, the enthalpy change reaches a plateau and remains almost constant up to 1000 ns (Figure 3). Hence, a 10 ns simulation is sufficient to reliably evaluate internal energies of oligonucleotides and their complexes and the hybridization enthalpy change on the basis of the MD trajectory.

5. Ionic Strength in the Implicit Water Shell. The influence of ionic strength on the structure and energy values was studied by the simulation of DDD in the concentration range of monovalent cations from 0.01 to 1 M in implicit water shell and the Andersen thermostat at 300 K.[30] The obtained structures of oligonucleotide complexes averaged along the 10 ns trajectory were almost equal (see $RMSD_{1ds}$ values in Table 1, stage 5), and the energy values calculated by the MD method somewhat decreased with the increase of the ionic strength. The decrease of the energy of dsDNA helix was more pronounced as compared with the single-stranded state, which led to the slight

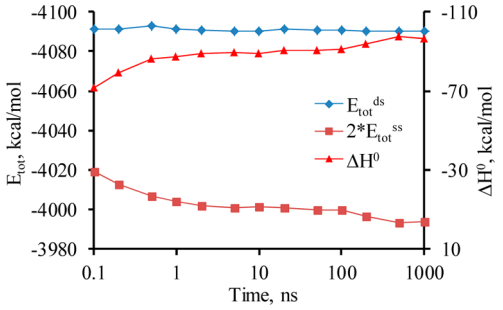


Figure 3. Dependence of full energy of single-stranded ($2*E_{tot}^{ss}$) and double-stranded (E_{tot}^{ds}) states (left axis) and the enthalpy of the self-complementary dodecamer complex formation ($\Delta H^\circ \approx E_{tot}^{ds} - 2*E_{tot}^{ss}$) (right axis) on simulation time in the implicit water shell (Andersen thermostat, 300 K, 0.1 M 1:1 salt) after heating of molecular dynamic system.

decrease in the hybridization enthalpy with the rise of the ionic strength. Variations of the enthalpy values of the complex formation at different ionic strength were ∼10%, which is comparable to the experimental uncertainty of the evaluation of the thermal stability of DNA duplexes. As the force field for the implicit water shell was parametrized for the buffer solutions containing ∼100 mM monovalent cations,[30] we have chosen these specific conditions.

6. Temperature Control Method. The influence of the temperature control method on the results of MD simulation was analyzed using Andersen or Langevin thermostats in implicit solvent shell. In the latter case, the recommended values of collision frequency are 2−5 ps^{-1}, but we have extended this range from 1 to 50 ps^{-1}. The value of 50 ps^{-1} corresponds to the collision frequency in aqueous solutions at room temperature.[57] It was found that the value of hybridization enthalpy change decreases with the rise of the collision frequency (Table 1) and, on average, is close to that obtained in the Andersen thermostat. It is worth noting that both types of thermostats are characterized by some artifacts. In the case of Langevine thermostat, there are "synchronization" artifacts and a quick denaturation of the biomolecule.[58,59] In addition, the usage of Langevine thermostat slightly decreases the calculation performance.[37] Taking this into account, we have chosen Andersen temperature coupling scheme for further simulations.

7. Water Shell Radius. The influence of the water shell radius on the structure and thermodynamic characteristics of oligonucleotide complexes in the case of the explicit solvent was studied. A cuboid box with the closest distance between any atoms originally presents in solute; the edge of the periodic box was in the range of 6−16 Å, and the TIP3-water model was applied. The results show that the structures of DNA duplexes averaged over the MD trajectories both in the implicit and in the explicit water shells are close to each other, independent of the water shell size ($RMSD_{1ds}$ < 2 Å, Table 1). In contrast, the total internal energy values of all simulated systems (oligonucleotide complexes in water) strongly depend on the radius of the water shell due to the different number of water molecules. To obtain the hybridization enthalpy, we used the molecular mechanics generalized Born (or Poisson−Boltzmann) surface area (MMGB(PB)SA) calculations. These methods imply the removal of water molecules and cations from the MD system, and the subsequent calculation of the internal energy values of the DNA states in the implicit water shell. In this case, either the total energies or the hybridization

DOI: 10.1021/acs.jpcb.5b09645
J. Phys. Chem. B 2015, 119, 15221−15234

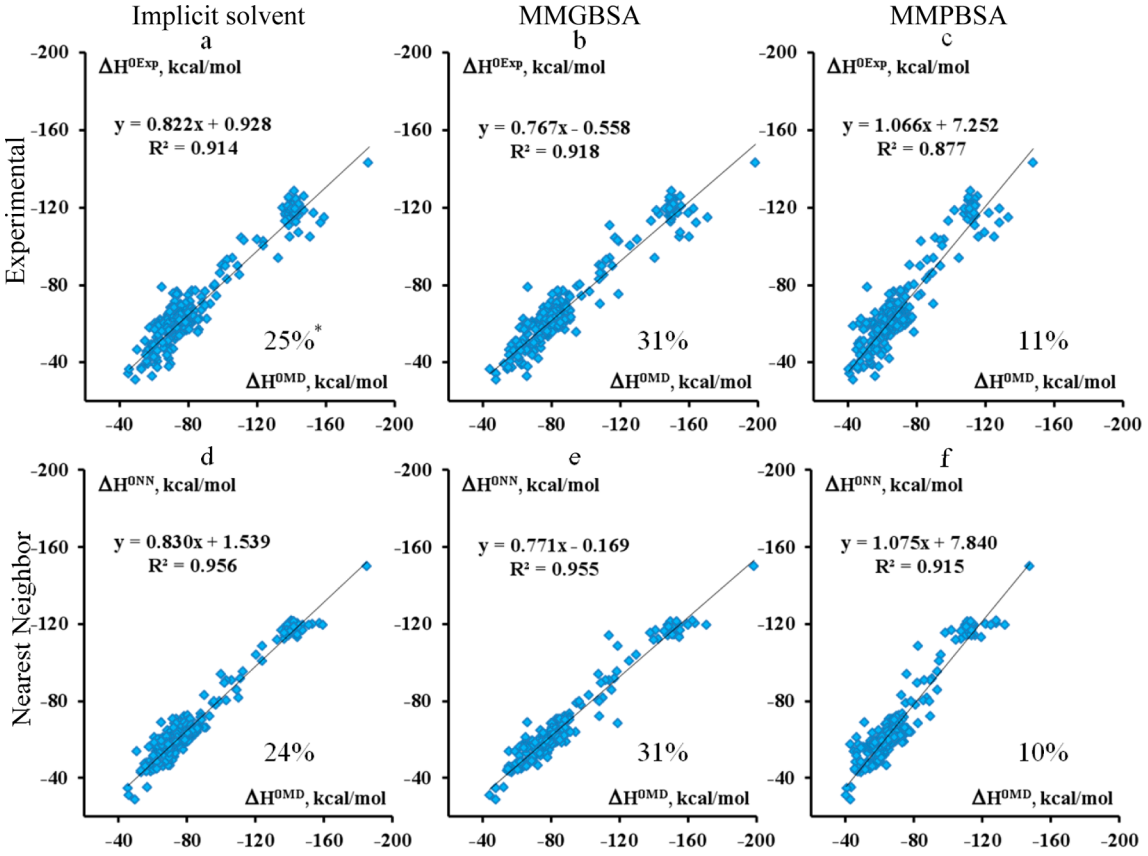


Figure 4. Correlation of the enthalpy values obtained in experiments (a, b, c) or calculated using the NN model (d, e, f) with those calculated by the MD method using optimal simulation parameters in the implicit (a, d) and explicit water shells [MMGBSA (b, e), MMPBSA (c, f)]. Asterisk indicates mean absolute values of the calculation errors.

enthalpy does not depend on radius of water shell. Furthermore, internal energies of DNA duplex formation calculated with the use of MMPB(GB)SA are comparable with those obtained in the implicit water shell where ssDNA and dsDNA were simulated (Table 1). These results are close to experimental data, as well as those calculated by the nearest neighbor approximation (the difference between the hybridization enthalpy values is ~20%). We have chosen 12 Å radius as the optimal value which allows enhancing the performance of MD simulations.

In accordance with the data obtained, the optimal simulation parameters for the implicit solvent shell include the following: the temperature 300 K using the Andersen temperature coupling scheme, SHAKE algorithm for bonds involving hydrogen, the time step of 1 fs, and the removing of translational and rotation motion of the center of mass every 1 ps. Productive simulation time in the implicit water shell should be 10 ns after heating to the 300 K for 2.5 and 1 ns equilibration. These times were selected as short enough for obtaining equilibrium MD trajectory.

In the case of the explicit water shell, after the equilibration of nucleic acids in the implicit solvent (as described above), sodium ions were added to the system to neutralize the net charge, together with water molecules. The TIP3 water model with 12 Å radius was used. The particle mesh Ewald (PME) method with 10 Å cutoff was applied. The length of productive MD trajectory was 10 ns.

The simulation in the explicit water shell should be performed in the TIP3P box (12 Å radius) after equilibrating for 10 ns, and the MMPB(GB)SA method should be applied to calculate the energy of the intermolecular complex formation. Details of the protocols are provided in the Materials and Methods section.

MD Simulation of a Set of Oligonucleotide Complexes. To prove the feasibility of the MD method for the reliable prediction of thermodynamic parameters of DNA duplex of any given sequence, it was necessary to calculate these parameters for the set of oligonucleotides of various lengths and GC contents. For this purpose, we have compiled a database of enthalpy, entropy, and Gibbs free energy changes for 305 native fully complementary intermolecular oligodeoxyribonucleotide complexes with the length from 4 to 16 bp and the GC content from 0 to 100% (Supporting Information, Table S1). The thermal denaturation of the most oligonucleotide complexes was fulfilled in accordance with the two-state model, which means that single-stranded oligonucleotides do not form hairpin structures or intermolecular complexes except the bimolecular fully complementary duplex.[60]

The energy of the formation of each complex was calculated by the MD simulation in explicit and implicit water shells using previously determined optimal parameters for DDD. The simulation of two single-stranded oligonucleotides and their complexes was carried out in the implicit water shell. In the explicit water shell, only the complexes of oligonucleotides were simulated. MD trajectories of the oligonucleotide complexes in the explicit water shell were treated using MMPBSA and the MMGBSA calculations. It was found that the enthalpy values of the complex formation calculated by the molecular dynamics

DOI: 10.1021/acs.jpcb.5b09645
J. Phys. Chem. B 2015, 119, 15221−15234

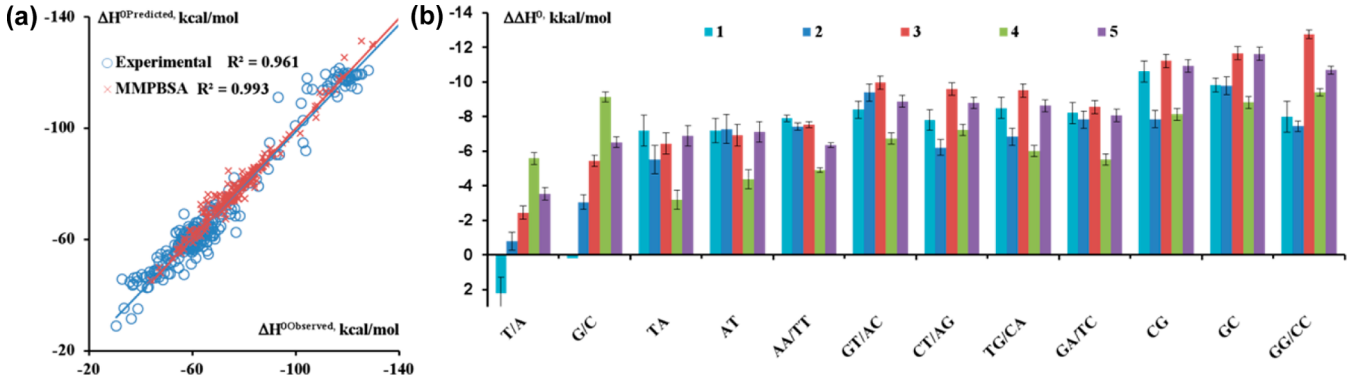


Figure 5. (a) Correlation of the hybridization enthalpy obtained experimentally or calculated by the MD method in the explicit water shell for 10 ns using the PB model with values calculated by the NN approximation using corresponding increments for each method. (b) The comparison of thermodynamic increments of the NN approximation: 1, experimental data from ref 1; 2, results of the analysis of experimental database; and results of the analysis of values calculated by the MD method, 3, explicit solvent, MMGBSA; 4, explicit solvent, MMPBSA; 5, implicit solvent.

method correlated very well with those obtained in experiments and calculated using the nearest neighbor model[3] (Figure 4).

The comparison of the enthalpy values obtained by the nearest neighbor approximation in 1 M NaCl with those calculated by the MD method in 0.1 M 1:1 salt is justified because experimental enthalpy values do not change in the range of monovalent cations concentration 0.01−1 M.[61−63] Figure 4 shows the mean absolute values of the calculation of hybridization enthalpy values. It can be seen that the simulation using the explicit water shell and MMPBSA ensures more reliable prediction of the enthalpy of DNA complex formation. The error is ∼10% in the case of the Poisson−Boltzmann and ∼30% when using the generalized Born model. Nevertheless, the coefficient of the correlation between experimental and calculated values of the hybridization enthalpy evaluated by the GB model is higher than in the case of the PB model (the R^2 values of the correlation with experimental values are 0.92 and 0.88, respectively). This is due to the greater slope of the trend lines (Figure 4b,c). In the first case, the slope 0.767 and the intercept close to zero lead to the overestimation of the enthalpy values for MMPBSA calculation. In the second case, the slope is close to 1, but the intercept is 7.25 kcal/mol. The GB model has a high absolute error, but low scatter about the regression line.

In all cases, the correlations between the enthalpy change values calculated by MD and nearest neighbor methods are better than the correlations between the values calculated by the MD method and experimental ones. This is probably caused by the fact that the NN model implies an ideal process of the complex formation corresponding to the two-state model. Random errors can occur when the internal hybridization energy is evaluated in experiments. Besides, single-stranded oligonucleotides can form ordered intramolecular structures, which have an impact on the calculated values of the hybridization energy.[64,65] However, in our calculations by MD methods, the effect of intramolecular ordering of the single-stranded oligonucleotide structures was not taken into account. The difference between thermodynamic parameters obtained experimentally and calculated by the MD method can also be attributed to both the faulty accounting of the contribution of single-stranded oligonucleotides during the simulation, and incorrect parameters of the force field. In order to determine whether the molecular dynamics method can be applied for the calculation of the formation energy of individual structural elements of DNA duplexes (dinucleotide base pairs and the terminal base pairs, which require corrections), hybridization energy values were analyzed using the nearest neighbor approximation, as described in ref 1.

The analysis of experimental enthalpy values by this model and the use of the multiple linear regression technique show that almost all thermodynamic increments are statistically significant (typically p-level < 0.05, Supporting Information, Table S2). The increment associated with the formation of the terminal A/T pair, which was evaluated for the experimental values of the hybridization enthalpy, was the only exception. Its low statistical significance can be explained by the low energy contribution to the thermal stability of extended duplexes.

Figure 5a demonstrates the typical correlation of the enthalpy values for the DNA duplex formation calculated by the MD method in the explicit water shell using the MMPBSA model or the experimental values with the values obtained by the NN model. It can be seen that the nearest neighbor model describes the data evaluated by the MD method better than the experimental values (R^2 = 0.99 and 0.96, respectively). Figure 5b shows the thermodynamic increment values corresponding to the formation of structural elements of DNA duplexes, which were evaluated by the NN approximation, in comparison with those widely used to calculate the thermal stability of DNA duplexes.[1]

The results show that the values of the increments from ref 1 are similar to those based on the experimental data included in the database. However, the increments corresponding to the formation of the terminal base pairs evaluated from both experimental data and MD calculations are negative, in contrast to the previously determined parameters.[1] The values of the 10 other increments corresponding to the internal base pairs formation are in good correlation with previously published data. In all cases, the well-known trend of the dependence of thermodynamic increments on the nucleotide composition is preserved. The formation of GC-rich dinucleotide pairs is, on average, energetically more favorable than that of AT-rich pairs and depends on the nucleotide sequence in each particular step.

The comparison of the hybridization enthalpy values obtained experimentally and calculated by the MD method shows that the values of thermodynamic increments for the structural elements containing only A/T-pairs coincide within the error of determination, when simulating DNA complexes in the explicit water shell and analyzing the structures by the GB model. The absolute energy values of the formation of the dinucleotide steps containing G/C pairs are rather over-

DOI: 10.1021/acs.jpcb.5b09645
J. Phys. Chem. B 2015, 119, 15221−15234

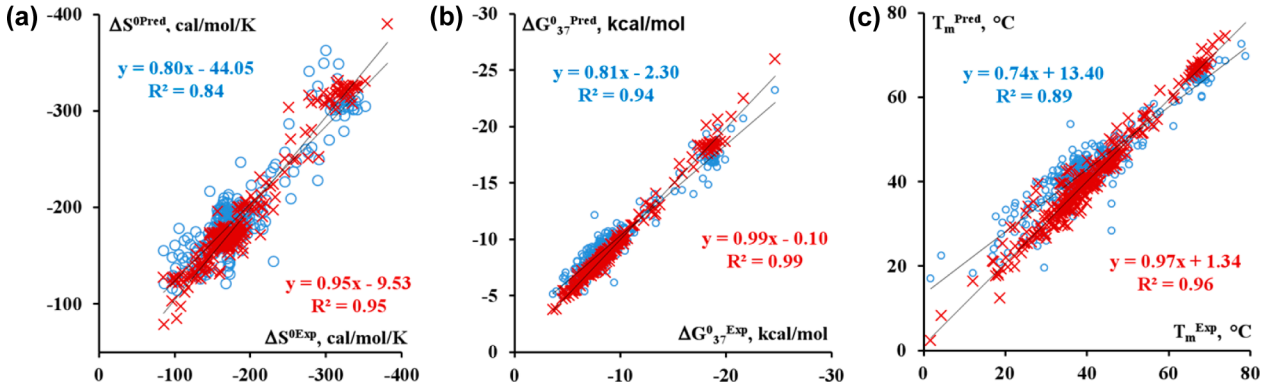


Figure 6. Correlation of the values predicted within NN model (red ×) and based on MD simulation in the explicit water using MMPBSA and eq 3 (blue ○) with experimental data: (a) entropy change values, (b) free Gibbs energy changes, (c) melting temperatures. Lower equation and R^2 values correspond to the NN model, and the upper corresponds to the values calculated on the basis of MD simulation data.

estimated. The dinucleotide base pair 5′-GG-3′/5′-CC-3′ gives the maximum contribution to the thermal stability of DNA duplexes (−12.7 kcal/mol).

The contributions of the individual dinucleotide steps to the hybridization energy were analyzed by the simulation of DNA duplexes in the explicit water shell with the treatment of the results using the PB model. It was revealed that A/T- and G/C-pairs stabilize DNA duplexes to a lesser or the same extent, respectively, in comparison with the thermodynamic increments derived from the experimental database.

In the case of the implicit water shell, the contribution of A/T- and G/C-containing dinucleotides to the thermal stability of DNA duplexes coincides with or exceeds the experimental values within the experimental uncertainty, respectively.

Thus, by the example of the representative database containing the thermal stability data for 305 oligonucleotide complexes of various GC contents and lengths, it has been shown that the method of molecular dynamics allowed the calculating of DNA duplex formation enthalpy with a high accuracy. The use of the ff99bsc0 force field, the explicit water shell, and the MMPBSA method for the calculation of hybridization energy ensures the prediction of the enthalpy change for duplexes with the accuracy of about 10%.

II. Evaluation of the Gibbs Free Energy Changes and the Melting Temperatures of the DNA/DNA Duplex Formation. One of the major tasks in the predictive calculation of the thermal stability of nucleic acid complexes is the evaluation of the Gibbs free energy changes and melting temperatures of DNA duplexes. The Gibbs free energy changes could be calculated using enthalpy and entropy values according to eq 2:

$$\Delta G^\circ(T) = (\Delta H^\circ - T\Delta S^\circ)\ \text{cal mol}^{-1} \quad (2)$$

As we have shown, the enthalpy values could be calculated by means of MD simulations. The entropy change values for the native DNA duplex formation could be calculated on the basis of the well-known linear correlation between the experimental enthalpy and entropy values of the complex formation for native duplexes.[65−67] The possible origin of this $\Delta H^\circ/\Delta S^\circ$ compensation was discussed earlier.[67] The $\Delta H^\circ/\Delta S^\circ$ correlation for the oligonucleotides included in our database is presented in Supporting Information Figure S4. There is a linear dependence of the hybridization entropy on the enthalpy of the complex formation with a very high correlation coefficient (R^2 = 0.995). This dependence is described by the equation

$$\Delta S^\circ = 2.678\Delta H^\circ/1000 - 6.0\ \text{cal mol}^{-1}\ \text{K}^{-1} \quad (3)$$

The experimental correlation is shown at 1 M NaCl concentration. Notably, this correlation can be considered as suitable for the calculation of the entropy using the enthalpy value evaluated by the MD method since the hybridization enthalpy does not depend significantly on the solution's ionic strength in a wide concentration range of cations.[61,62] The usage of eq 3 for the calculation of the hybridization entropy on the basis of the enthalpy change values determined by the MD method (Figure 6a) gave the mean absolute value of the error of 11.6%, which is comparable to the accuracy of ΔS° evaluation by the experimental approaches[68] and the calculation by the nearest neighbor model (6%, Figure 6a). This finding suggests that the method might be used to calculate the full set of thermodynamic parameters of DNA duplex formation.

The hybridization enthalpy values obtained by MD method and the entropy values calculated according to eq 3 were used for the calculations of the Gibbs energy values (at 37 °C), which were close to the experimental values, with accuracy of 10% (Figure 6b). At the same time, the average error of the free energy prediction using the nearest neighbor model was 3.1% for all duplexes in the database (Figure 6b).

The melting temperatures (T_m's) for DNA/DNA complexes were calculated using eq 4

$$T_m = (\Delta H^\circ/(\Delta S^\circ + R\ \ln(C_T/\gamma\ )))\text{K} \quad (4)$$

where R = 1.987 cal/mol/K (gas constant), C_T is the total concentration of a stoichiometric mixture of oligonucleotides in solution, γ = 4 for a non-self-complementary complex, and γ = 2 for a self-complementary complex.[69] The thermal stability was calculated for 10^{-5} M total concentration of the oligonucleotides.

It was shown that the use of the simulation data in the explicit solvent, together with MMPBSA and eq 3, for the calculation of the hybridization enthalpy allows evaluating the melting temperature with the greatest accuracy (Supporting Information, Figure S5). The mean absolute value of the calculation error was 4.4 °C while its standard deviation was 3.3 °C. The correlations of experimental and calculated values with MD simulation thermodynamic parameters and melting temperatures are presented in Figure 6c. The error value for the calculation of T_m by the computer simulation significantly

DOI: 10.1021/acs.jpcb.5b09645
J. Phys. Chem. B 2015, 119, 15221−15234

exceeds the typical experimental values (0.5 °C) and is only 3 times higher than the error for the nearest neighbor method (1.3 °C, Figure 6c). Nevertheless, our study shows that the molecular dynamics method gives a surprisingly good prediction of the thermal stability of the DNA duplexes of various length and nucleotide composition.

DISCUSSION

A predictive calculation of the thermal stability for the complexes of nucleic acids, their analogues, and derivatives opens broad prospects for a variety of technologies, applications, and tools in engineering and basic science. Methods known to date, including the most commonly used nearest neighbor approach, provide accurate predictions of the thermodynamic parameters only for nucleic acid complexes which are structurally close to previously examined ones. These include the native DNA/DNA, DNA/RNA, and RNA/RNA complexes, either perfectly matched or containing mismatches, bulges, and loops.[1,65,70] The predictive calculations of the thermal stability of modified duplexes are possible only for a limited set of derivatives and analogues such as deoxyinosine containing oligomers,[71] locked nucleic acids (LNA),[2] peptide nucleic acids (PNA),[72,73] non-nucleotide insertions of oligo-ethylene glycol phosphodiesters,[3,33] and a few others. The commonly used approaches do not allow accurate predictive calculations for the complexes containing the noncharacterized modifications.[74] The molecular dynamics method is devoid of such limitation, and so it is suitable for characterizing the structure, conformational mobility, and energy of new nucleic acid derivatives and analogues. However, the quantitative characterization of oligonucleotide duplexes of various length and nucleotide composition by the MD method has not yet been demonstrated. A few works have been published showing the possibility of predictive calculations of the thermal stability for a small set of oligonucleotide complexes. For example, the Gibbs free energy has been calculated for the formation of the $d(CGCG)_2$ complex,[22] and the melting temperature has been calculated by the MD simulation for DDD including the case of high cation concentrations.[23] It has also been found that melting temperatures for a number of modified complexes can be predicted from the Gibbs free energy values calculated by the MD method.[24] We should, however, point to a significant simplification of the approach in the latter publication.[26] The authors relied on the correlation of observed values of the Gibbs free energy change with the melting temperature at certain experimental conditions. This correlation cannot form a basis for the accurate predictive calculations of the thermal stability of nucleic acid complexes due to the dependence of T_m on the oligonucleotide concentrations, according to eq 4. To evaluate the advantages and limitations of *in silico* calculations, it would also be interesting to analyze the correlation of the values of the Gibbs free energy obtained experimentally with those calculated by the MD simulation. Nevertheless, the authors of ref 24 have shown the potential of the molecular dynamics for indirect calculation of the thermal stability for a number of native and modified oligonucleotide complexes.

To determine the thermal stability of the duplex via the MD simulation, it is necessary to calculate the total internal energies of single- and double-stranded oligonucleotides. We have studied the influence of the simulation parameters on the evaluated energy values. The optimal parameters for the calculation of hybridization enthalpy have been determined in cases of the implicit and explicit water shell. We have found that the parameters obtained are very close to those that are widely used in the studies of structure and conformational mobility of nucleic acids (e.g., see ref 27).

Moreover, typical energy values for DDD duplex in the explicit solvent calculated with the MMPBSA approximation (about −3000 kcal/mol) are close to the values obtained using of the CHARMM c32b1 molecular mechanics package, the CHARMM27 all-atom NA force field, and the MMPBSA method (∼ −3200 kcal/mol).[75] This is a surprisingly good correlation, since the energy of double-stranded DDD varies in a wide range (from −2900 up to −5500 kcal/mol, Table 1) depending on the simulation parameters.

Another important issue concerns the applicability of the force field to the single-stranded DNA. The origin of the problem is that preliminary developed force field parameters are based on the data for double-stranded nucleic acids. The analysis of simulation of single-stranded oligonucleotides shows that it is necessary to add weak harmonic potential to prevent compactization. The small values of restraint potential together with the same types of interactions and close energy values of single- and double-stranded oligonucleotides could probably cause the compensation effect for the calculation of the enthalpy changes. This suggestion can be proved by much better correlation of the experimental data with those obtained from the simulations in explicit solvent. In this case, MD trajectories of single-stranded states are derived from the double-stranded, so the conformations are highly consistent. This leads to the high extent of correlation between the total energies of single- and double-stranded states and thus the low hybridization enthalpy variation (−91.5 to −129.1 kcal/mol). These data are in very good agreement with the experimental value (−116 kcal/mol). This fact in combination with a good agreement between experimental and calculated dsDNA spatial structure indicates a sufficient parametrization of the force field for the duplexes.

In this paper, we have shown for the first time the possibility of using MD simulation for the high accuracy calculation of the hybridization enthalpy as a part of the free energy by an example of the large number (305) of the DNA duplexes. We have compiled the representative database of thermodynamic parameters for oligonucleotide hybridization. The length and GC-content, and hence the formation energies and melting temperatures of nucleic acid complexes, vary in a wide range (Supporting Information, Table S1). Consequently, the obtained results are statistically significant, and can be applied to other native DNA oligonucleotide complexes beyond this database.

Comparative studies of the MMPBSA and MMGBSA approximations for the calculation of hybridization energy show that in the first case the enthalpy change values are calculated more accurately (average absolute error of ΔH° = 10.6% and 30.6%, respectively), but the correlation with experimental values is higher in the second case (R^2 = 0.88 and 0.92, respectively). Similar results have been observed for these two models previously in the studies of ligand−protein interactions.[76] It has been shown that MMGBSA gives worse predictions for the absolute binding free energies than MMPBSA, but shows more correct values for ranking of the ligand binding affinities. Similar observations were obtained for calculations in the case of the implicit solvent water shell. The value of R^2 was very high (0.91), but the accuracy of the enthalpy change prediction was low (absolute error is 25%) (Figure S5d), and could probably be the result of using the

DOI: 10.1021/acs.jpcb.5b09645
J. Phys. Chem. B 2015, 119, 15221−15234

generalized Born approximation instead of the exact Poisson−Boltzmann equation.[77]

The prediction ability of molecular dynamic simulation for calculating an enthalpy change value is close to that of the nearest neighbor model which is now considered to be the best one. The absolute error values in the second case are 5%, only twice lower than in the case of computer simulation in implicit water. These results, together with MD simulation data for nucleic acid derivatives, and the same general principles of interaction for nucleic acids derivatives and native DNA and RNA, could probably give the opportunity for efficient predictive calculation of hybridization enthalpy for the nucleic acids analogues. Of course, to obtain the high precision of such calculations, some improvement is obligatory, such as force field optimization for single-stranded nucleic acids, force field testing for a number of derivatives, etc. Therefore, proposed methods for DNA thermodynamics calculations are not as much an alternative to the nearest neighbor model as the background for future developments.

To quantitatively characterize the thermal stability of nucleic acid complexes, it is more appropriate to use the values of the Gibbs free energy change and melting temperature than ΔH°. Hence, the entropy change needs to be estimated to calculate ΔG°_{37} and T_m values. The computation of the ΔS° is not a trivial task because of the necessity to evaluate the entropy of nucleic acid interaction with solute and ions, that was demonstrated experimentally.[62,78−80] For the evaluation of the hybridization entropy, we propose the method relying on the phenomenological $\Delta H^\circ/\Delta S^\circ$ correlation as a good approximation for high accuracy calculation of thermal stability of NA duplexes under certain buffer conditions. Moreover, such a type of correlations could be observed for nucleic acids containing mismatches, deoxyinosine, and modifications which are widely studied: LNA and non-nucleotide oligoethelene glycol inserts (see Supporting Information Figure S6). Thus, linear correlation could probably be observed for the newly developed modifications, because the general principles of NA complex formation are proposed to remain the same. At the same time using entropy/enthalpy compensation formula for prediction of new derivatives may be risky because there is no *a priori* method to observe such type of correlation. So, the calculation of entropy or the free Gibbs energy changes is too complicated a problem and will be the subject of our future studies.

The results obtained in ref 24, for modified oligonucleotides, show worse correlation between the calculated and measured melting temperatures (R^2 = 0.84 in the range of experimental T_m 60−90 °C) than our data (R^2 > 0.87 for various methods of calculation in the range of experimental T_m 2−79 °C). Actually, we observed the excellent accuracy of predictions of binding Gibbs free energies (10%) and melting temperatures (4.4 °C) (Figure 6). Despite quite good correlation of the melting temperatures, the systematic underestimating of calculated data was observed. We suppose two reasons for this effect: the dependence on the fraction of G/C pairs, and the duplex length. The absolute error value for the Gibbs energy change calculation is typical for the NN method as well as for the experimental determination. The error for the melting temperature calculation is only 3 times higher than typical values in the case of prognostic NN model-based calculation.

The NN model analysis of the hybridization enthalpies calculated by MD approach shows that MD-simulated data fit the NN model very well (Figure 5a). The correlation in the case of $\Delta H^{\circ MD}$ is much higher than in the case of $\Delta H^{\circ Exp}$ (R^2 = 0.993 and 0.961). In addition, the values of thermodynamic increments corresponding to the formation of various dinucleotide pairs are symbatic (Figure 5b). It is well-known that the dinucleotide pair formation is more favorable in the case of GC-containing helix steps. Similar results have been observed for all simulation methods. Moreover, the thermodynamic increments are close to each other in the case of generalized Born approximation in explicit and implicit solvation models. The obtained results could be taken into account for optimization of the force field parameters.

CONCLUSIONS

In this paper we have shown for the first time that the enthalpy of the formation of complementary DNA complexes can be calculated with good accuracy by the molecular dynamics simulation. The impacts of the explicit and implicit water shells, temperature, heating time, thermostat model, radius of the water shell, and simulation time were studied on the example of the Dickerson−Drew dodecamer. We have determined the optimal parameters of a computer experiment, which resulted in the hybridization enthalpy value very close to the experimentally obtained one. To estimate the reliability of prediction of the enthalpy values, we created the database containing experimentally obtained thermodynamic parameters for the formation of DNA duplexes. The simulation of a set of 305 oligonucleotide complexes was carried out for the explicit and implicit water shells with various parameters. It has been shown that the enthalpy calculation with a maximum achievable accuracy requires the 10 ns simulation in the ff99bsc0 force field, the explicit water shell, and the consequent analysis of the complex using the MMPBSA calculation. In this case, a mean error of the enthalpy calculation for the formation of DNA duplexes was 10.6%. The hybridization enthalpy evaluated by the MD simulation and linear correlation between experimental enthalpy and entropy of the duplex formation have been used to calculate the Gibbs free energy and melting temperature for all duplexes. It has been found that the proposed simulation method can be used for the *a priori* calculation of the Gibbs free energy of oligonucleotide hybridization and the melting temperature of corresponding duplexes with an accuracy of 10% and 4.4 °C, respectively.

This work is the first step to predictive calculation of the thermal stability of nucleic acids, their derivatives, and analogues with the use of molecular dynamics simulation. In the subsequent communication, we are going to apply the developed methodology to the predictive calculation of the thermal stability of modified oligonucleotide complexes.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.5b09645.

Additional thermodynamic parameters, calculations, and structures (PDF)

AUTHOR INFORMATION

Corresponding Authors

*Phone: (+7)-383-365-5134. Fax: (+7)-383-365-5153. E-mail: lomzov@niboch.nsc.ru.

*Phone: (+7)-383-365-5151. Fax: (+7)-383-365-5153. E-mail: pyshnyi@niboch.nsc.ru.

DOI: 10.1021/acs.jpcb.5b09645
J. Phys. Chem. B 2015, 119, 15221−15234

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Prof. S. Altman for carefully reading the manuscript and providing helpful suggestions. Sincere gratitude is extended to Prof. Don Crothers for his comments and suggestions. This work was done in accordance with State assignment and supported by the Russian Foundation for Basic Research [No. 13-04-01176], by the Russian Government support for research projects implemented under supervision of world's leading scientists with Sidney Altman as the leading scientist [No. 14.B25.31.0028], by an interdisciplinary grant from the Siberian Branch of Russian Academy of Sciences, and by a grant from the Russian Science Foundation (No. 14-14-00922) for part of the computational algorithms optimization. We greatly appreciate the Siberian Supercomputer Centre Siberian Branch RAS for providing the computational resource.

■ REFERENCES

(1) SantaLucia, J.; Hicks, D. The Thermodynamics of DNA Structural Motifs. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 415−440.

(2) McTigue, P. M.; Peterson, R. J.; Kahn, J. D. Sequence-Dependent Thermodynamic Parameters; for Locked Nucleic Acid (LNA)-DNA Duplex Formation. *Biochemistry* **2004**, *43*, 5388−5405.

(3) Lomzov, A. A.; Pyshnaya, I. A.; Ivanova, E. M.; Pyshnyi, D. V. Thermodynamic Parameters for Calculating the Stability of Complexes of Bridged Oligonucleotides. *Dokl. Biochem. Biophys.* **2006**, *409*, 211−215.

(4) Haran, T. E.; Mohanty, U. The Unique Structure of A-Tracts and Intrinsic DNA Bending. *Q. Rev. Biophys.* **2009**, *42*, 41−81.

(5) Kastenholz, M. A.; Schwartz, T. U.; Hunenberger, P. H. The Transition Between the B and Z Conformations of DNA Investigated by Targeted Molecular Dynamics Simulations with Explicit Solvation. *Biophys. J.* **2006**, *91*, 2976−2990.

(6) Il'icheva, I. A.; Vlasov, P. K.; Esipova, N. G.; Tumanyan, V. G. The Intramolecular Impact to the Sequence Specificity of B- > A Transition: Low Energy Conformational Variations in AA/TT and GG/CC Steps. *J. Biomol. Struct. Dyn.* **2010**, *27*, 677−693.

(7) Beveridge, D. L.; Barreiro, G.; Byun, K. S.; Case, D. A.; Cheatham, T. E.; Dixit, S. B.; Giudice, E.; Lankas, F.; Lavery, R.; Maddocks, J. H.; et al. Molecular Dynamics Simulations of the 136 Unique Tetranucleotide Squences of DNA Oligonucleotides. I. Research Design and Results on d(C(p)G) Steps. *Biophys. J.* **2004**, *87*, 3799−3813.

(8) Dixit, S. B.; Beveridge, D. L.; Case, D. A.; Cheatham, T. E.; Giudice, E.; Lankas, F.; Lavery, R.; Maddocks, J. H.; Osman, R.; Sklenar, H.; et al. Molecular Dynamics Simulations of the 136 Unique Tetranucleotide Sequences of DNA Oligonucleotides. II: Sequence Context Effects on the Dynamical Structures of the 10 Unique Dinucleotide Steps. *Biophys. J.* **2005**, *89*, 3721−3740.

(9) Dixit, S. B.; Mezei, M.; Beveridge, D. L. Studies of Base Pair Sequence Effects on DNA Solvation Based on All-Atom Molecular Dynamics Simulations. *J. Biosci.* **2012**, *37*, 399−421.

(10) Egli, M. DNA-Cation Interactions: Quo Vadis? *Chem. Biol.* **2002**, *9*, 277−286.

(11) Yoo, J.; Aksimentiev, A. Competitive Binding of Cations to Duplex DNA Revealed through Molecular Dynamics Simulations. *J. Phys. Chem. B* **2012**, *116*, 12946−12954.

(12) Jayaram, B.; Beveridge, D. L. Modeling DNA in Aqueous Solutions: Theoretical and Computer Simulation Studies on the Ion Atmosphere of DNA. *Annu. Rev. Biophys. Biomol. Struct.* **1996**, *25*, 367−394.

(13) Treesuwan, W.; Wittayanarakul, K.; Anthony, N. G.; Huchet, G.; Alniss, H.; Hannongbua, S.; Khalaf, A. I.; Suckling, C. J.; Parkinson, J. A.; Mackay, S. P. A Detailed Binding Free Energy Study of 2:1 Ligand−DNA Complex Formation by Experiment and Simulation. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10682−10693.

(14) Pande, V.; Nilsson, L. Insights into Structure, Dynamics and Hydration of Locked Nucleic Acid (LNA) Strand-Based Duplexes from Molecular Dynamics Simulations. *Nucleic Acids Res.* **2008**, *36*, 1508−1516.

(15) Chaubey, A. K.; Dubey, K. D.; Ojha, R. P. Stability and Free Energy Calculation of LNA Modified Quadruplex: a Molecular Dynamics Study. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 289−299.

(16) Rathinavelan, T.; Yathindra, N. Molecular Dynamics Structures of Peptide Nucleic Acid x DNA Hybrid in the Wild-Type and Mutated Alleles of Ki-ras Proto-Oncogene-Stereochemical Rationale for the Low Affinity of PNA in the Presence of an AC Mismatch. *FEBS J.* **2005**, *272*, 4055−4470.

(17) Blas, J. R.; Huertas, O.; Tabares, C.; Sumpter, B. G.; Fuentes-Cabrera, M.; Orozco, M.; Ordejon, P.; Luque, F. J. Structural, Dynamical, and Electronic Transport Properties of Modified DNA Duplexes Containing Size-Expanded Nucleobases. *J. Phys. Chem. A* **2011**, *115*, 11344−11354.

(18) Shevelev, G. Y.; Krumkacheva, O. A.; Lomzov, A. A.; Kuzhelev, A. A.; Rogozhnikova, O. Y.; Trukhin, D. V.; Troitskaya, T. I; Tormyshev, V. M.; Fedin, M. V.; Pyshnyi, D. V.; et al. Physiological-Temperature Distance Measurement in Nucleic Acid Using Triarylmethyl-Based Spin Labels and Pulsed Dipolar EPR Spectroscopy. *J. Am. Chem. Soc.* **2014**, *136*, 9874−9877.

(19) Beveridge, D. L.; Cheatham, T. E., III; Mezei, M. The ABCs of Molecular Dynamics Simulations on B-DNA, Circa 2012. *J. Biosci.* **2012**, *37*, 379−397.

(20) Ricci, C. G.; de Andrade, A. S.; Mottin, M.; Netz, P. A. Molecular Dynamics of DNA: Comparison of Force Fields and Terminal Nucleotide Definitions. *J. Phys. Chem. B* **2010**, *114*, 9882−9893.

(21) Pérez, A.; Luque, F. J.; Orozco, M. Frontiers in Molecular Dynamics Simulations of DNA. *Acc. Chem. Res.* **2012**, *45*, 196−205.

(22) Kannan, S.; Zacharias, M. Simulation of DNA Double-Strand Dissociation and Formation During Replica-Exchange Molecular Dynamics Simulations. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10589−10595.

(23) Izanloo, C.; Parsafar, G. A.; Abroshan, H.; Akbarzadeh, H. Denaturation of Drew-Dickerson DNA in a High Salt Concentration Medium: Molecular Dynamics Simulations. *J. Comput. Chem.* **2011**, *32*, 3354−3361.

(24) Shen, L. L.; Johnson, T. L.; Clugston, S.; Huang, H. W.; Butenhof, K. J.; Stanton, R. V. Molecular Dynamics Simulation and Binding Energy Calculation for Estimation of Oligonucleotide Duplex Thermostability in RNA-Based Therapeutics. *J. Chem. Inf. Model.* **2011**, *51*, 1957−1965.

(25) Yildirim, I.; Kennedy, S. D.; Stern, H. A.; Hart, J. M.; Kierzek, R.; Turner, D. H. Revision of AMBER Torsional Parameters for RNA Improves Free Energy Predictions for Tetramer Duplexes with GC and iGiC Base Pairs. *J. Chem. Theory Comput.* **2012**, *8*, 172−181.

(26) Spasic, A.; Serafini, J.; Mathews, D. H. The Amber ff99 Force Field Predicts Relative Free Energy Changes for RNA Helix Formation. *J. Chem. Theory Comput.* **2012**, *8*, 2497−2505.

(27) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. An Overview of the Amber Biomolecular Simulation Package. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2013**, *3*, 198−210.

(28) Zgarbová, M.; Otyepka, M.; Šponer, J.; Mládek, A.; Banáš, P.; Cheatham, T. E., III; Jurecka, 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.

(29) Onufriev, A.; Bashford, D.; Case, D. A. Exploring Protein Native States and Large-Scale Conformational Changes with a Modified Generalized Born Model. *Proteins: Struct., Funct., Genet.* **2004**, *55*, 383−394.

(30) Srinivasan, J.; Trevathan, M. W.; Beroza, P.; Case, D. A. Application of a Pairwise Generalized Born Model to Proteins and

DOI: 10.1021/acs.jpcb.5b09645
J. Phys. Chem. B 2015, 119, 15221−15234

Nucleic Acids: Inclusion of Salt Effects. *Theor. Chem. Acc.* **1999**, *101*, 426−434.
(31) Berendsen, H. J.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A. R. H. J.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(32) Lokhov, S. G.; Pyshnyi, D. V. Thermodynamic and Spectral Properties of DNA Miniduplexes with the Terminal G x A Mispairs and 3′ or 5′ Dangling Bases. *FEBS Lett.* **1997**, *420*, 134−138.
(33) Pyshnyi, D. V.; Lomzov, A. A.; Pyshnaya, I. A.; Ivanova, E. M. Hybridization of the Bridged Oligonucleotides with DNA: Thermodynamic and Kinetic Studies. *J. Biomol. Struct. Dyn.* **2006**, *23*, 567−579.
(34) Allawi, H. T.; SantaLucia, J. Thermodynamics and NMR of Internal GT Mismatches in DNA. *Biochemistry* **1997**, *36*, 10581−10594.
(35) Doktycz, M. J.; Morris, M. D.; Dormady, S. J.; Beattie, K. L.; Jacobson, K. B. Optical Melting of 128-Octamer DNA Duplexes - Effects of Base-Pair Location and Nearest Neighbours on Thermal Stability. *J. Biol. Chem.* **1995**, *270*, 8439−8445.
(36) Tanaka, F.; Kameda, A.; Yamamoto, M.; Ohuchi, A. Thermodynamic Parameters Based on a Nearest-Neighbor Model for DNA Sequences with a Single-Bulge Loop. *Biochemistry* **2004**, *43*, 7143−7150.
(37) Case, D. A.; Darden, T. A.; Cheatham T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; et al. *AMBER 12*; University of California: San Francisco, CA, 2012.
(38) Drew, H. R.; Samson, S.; Dickerson, R. E. Structure of a B-DNA Dodecamer at 16K. *Proc. Natl. Acad. Sci. U. S. A.* **1982**, *79*, 4040−4044.
(39) Holbrook, S. R.; Dickerson, R. E.; Kim, S. H. Anisotropic Thermal-Parameter Refinement of the DNA Dodecamer CGCGAATTCGCG by the Segmented Rigid-Body Method. *Acta Crystallogr., Sect. B: Struct. Sci.* **1985**, *41*, 255−262.
(40) Drew, H. R.; Wing, R. M.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. Structure of a B-DNA Dodecamer - Conformation and Dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **1981**, *78*, 2179−2183.
(41) Denisov, A. Y.; Zamaratski, E. V.; Maltseva, T. V.; Sandström, A.; Bekiroglu, S.; Altmann, K. H.; Egli, M.; Chattopadhyaya, J. The Solution Conformation of a Carbocyclic Analog of the Dickerson-Drew Dodecamer: Comparison with its own X-ray Structure and that of the NMR Structure of the Native Counterpart. *J. Biomol. Struct. Dyn.* **1998**, *16*, 547−568.
(42) Dixit, S. B.; Ponomarev, S. Y.; Beveridge, D. L. Root Mean Square Deviation Probability Analysis of Molecular Dynamics Trajectories on DNA. *J. Chem. Inf. Model.* **2006**, *46*, 1084−1093.
(43) Pauwels, E.; Claeys, D.; Martins, J. C.; Waroquier, M.; Bifulco, G.; Van Speybroeck, V.; Madder, A. Accurate Prediction of 1H Chemical Shifts in Interstrand Cross-Linked DNA. *RSC Adv.* **2013**, *3*, 3925−3938.
(44) Shikiya, R.; Li, J. S.; Gold, B.; Marky, L. A. Incorporation of Cationic Chains in the Dickerson-Drew Dodecamer: Correlation of Energetics, Structure, and Ion and Water Binding. *Biochemistry* **2005**, *44*, 12582−12588.
(45) Timsit, Y. DNA-Directed Base Pair Opening. *Molecules* **2012**, *17*, 11947−11964.
(46) Yin, Y. D.; Zhao, X. S. Kinetics and Dynamics of DNA Hybridization. *Acc. Chem. Res.* **2011**, *44*, 1172−1181.
(47) Huang, N.; MacKerell, A. D. Atomistic View of Base Flipping in DNA. *Philos. Trans. R. Soc., A* **2004**, *362*, 1439−1460.
(48) Andreatta, D.; Sen, S.; Lustres, J. L. P.; Kovalenko, S. A.; Ernsting, N. P.; Murphy, C. J.; Coleman, R. S.; Berg, M. A. Ultrafast Dynamics in DNA: "Fraying" at the End of the Helix. *J. Am. Chem. Soc.* **2006**, *128*, 6885−6892.
(49) Zgarbova, M.; Otyepka, M.; Sponer, J.; Lankas, F.; Jurecka, P. Base Pair Fraying in Molecular Dynamics Simulations of DNA and RNA. *J. Chem. Theory Comput.* **2014**, *10*, 3177−3189.
(50) Sim, A. Y.; Lipfert, J.; Herschlag, D.; Doniach, S. Salt Dependence of the Radius of Gyration and Flexibility of Single-Stranded DNA in Solution Probed by Small-Angle X-ray Scattering. *Phys. Rev. E Stat. Nonlin. Soft Matter. Phys.* **2012**, *86*, 021901.
(51) Chen, H.; Meisburger, S. P.; Pabit, S. A.; Sutton, J. L.; Webb, W. W.; Pollack, L. Ionic Strength-Dependent Persistence Lengths of Single-Stranded RNA and DNA. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 799−804.
(52) Kang, J.; Jung, J.; Kim, S. K. Flexibility of Single-Stranded DNA Measured by Single-Molecule FRET. *Biophys. Chem.* **2014**, *195*, 49−52.
(53) Murphy, M. C.; Rasnik, I.; Cheng, W.; Lohman, T. M.; Ha, T. J. Probing Single-Stranded DNA Conformational Flexibility Using Fluorescence Spectroscopy. *Biophys. J.* **2004**, *86*, 2530−2537.
(54) Tinland, B.; Pluen, A.; Sturm, J.; Weill, G. Persistence Length of Single-Stranded DNA. *Macromolecules* **1997**, *30*, 5763−5765.
(55) Rechendorff, K.; Witz, G.; Adamcik, J.; Dietler, G. Persistence Length and Scaling Properties of Single-Stranded DNA Adsorbed on Modified. *J. Chem. Phys.* **2009**, *131*, 095103.
(56) Causley, G. C.; Staskus, P. W.; Johnson, W. C. Improved Methods of Analysis for CD Data Applied to Single-Strand Stacking. *Biopolymers* **1983**, *22*, 945−967.
(57) Pastor, R. W.; Brooks, B. R.; Szabo, A. An Analysis of the Accuracy of Langevin and Molecular-Dynamics Algorithms. *Mol. Phys.* **1988**, *65*, 1409−1419.
(58) Uberuaga, B.; Anghel, M.; Voter, A. Synchronization of Trajectories in Canonical Molecular-Dynamics Simulations: Observation, Explanation, and Exploitation. *J. Chem. Phys.* **2004**, *120*, 6363−6374.
(59) Sindhikara, D.; Kim, S.; Voter, A.; Roitberg, A. Bad Seeds Sprout Perilous Dynamics: Stochastic Thermostat Induced Trajectory Synchronization in Biomolecules. *J. Chem. Theory Comput.* **2009**, *5*, 1624−1631.
(60) Lomzov, A. A.; Pyshnyi, D. V. Considering the Oligonucleotide Secondary Structures in Thermodynamic and Kinetic Analysis of DNA Duplex Formation. *Biophysics* **2012**, *57*, 19−34.
(61) Oro, J. R. D.; Grigera, J. R. On the Thermal Stability of DNA in Solution of Mixed Solvents. *J. Biol. Phys.* **1995**, *21*, 151−154.
(62) Owczarzy, R.; You, Y.; Moreira, B. G.; Manthey, J. A.; Huang, L. Y.; Behlke, M. A.; Walder, J. A. Effects of Sodium Ions on DNA Duplex Oligomers: Improved Predictions of Melting Temperatures. *Biochemistry* **2004**, *43*, 3537−3554.
(63) Manning, G. S. The Molecular Theory of Polyelectrolyte Solutions with Applications to the Electrostatic Properties of Polynucleotides. *Q. Rev. Biophys.* **1978**, *11*, 179−246.
(64) Vesnaver, G.; Breslauer, K. J. The Contribution of DNA Single-Stranded Order to the Thermodynamics of Duplex Formation. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 3569−3573.
(65) Petruska, J.; Goodman, M. F.; Boosalis, M. S.; Sowers, L. C.; Cheong, C.; Tinoco, I. Comparison between DNA Melting Thermodynamics and DNA-Polymerase Fidelity. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85*, 6252−6256.
(66) Searle, M. S.; Williams, D. H. On the Stability of Nucleic Acid Structures in Solution: Enthalpy-Entropy Compensations, Internal Rotations and Reversibility. *Nucleic Acids Res.* **1993**, *21*, 2051−2056.
(67) Starikov, E. B.; Norden, B. Physical Rationale Behind the Nonlinear Enthalpy-Entropy Compensation in DNA Duplex Stability. *J. Phys. Chem. B* **2009**, *113*, 4698−4707.
(68) Xia, T. B.; SantaLucia, J.; Burkard, M. E.; Kierzek, R.; Schroeder, S. J.; Jiao, X. Q.; Cox, C.; Turner, D. H. Thermodynamic Parameters for an Expanded Nearest-Neighbor Model for Formation of RNA Duplexes with Watson-Crick Base Pairs. *Biochemistry* **1998**, *37*, 14719−14735.
(69) Breslauer, K. J. In *Thermodynamic Data for Biochemistry and Biotechnology*; Hinz, H.-J., Ed.; Springer: New York, 1986; pp 402−427.
(70) Wu, P.; Nakano, S.; Sugimoto, N. Temperature Dependence of Thermodynamic Properties for DNA/DNA and RNA/DNA Duplex Formation. *Eur. J. Biochem.* **2002**, *269*, 2821−2830.

DOI: 10.1021/acs.jpcb.5b09645
J. Phys. Chem. B 2015, 119, 15221−15234

(71) Watkins, N. E.; SantaLucia, J. Nearest-Neighbor Thermodynamics of Deoxyinosine Pairs in DNA Duplexes. *Nucleic Acids Res.* **2005**, *33*, 6258−6267.
(72) Giesen, U.; Kleider, W.; Berding, C.; Geiger, A.; Orum, H.; Nielsen, P. E. A Formula for Thermal Stability (Tm) Prediction of PNA/DNA Duplexes. *Nucleic Acids Res.* **1998**, *26*, 5004−5006.
(73) Takiya, T.; Seto, Y.; Yasuda, H.; Suzuki, T.; Kawai, K. An Empirical Approach for Thermal Stability (Tm) Prediction of PNA/DNA Duplexes. *Nucleic Acids Symp. Ser.* **2004**, *48*, 131−132.
(74) Freier, S. M.; Altmann, K. H. The Ups and Downs of Nucleic Acid Duplex Stability: Structure-Stability Studies on Chemically-Modified DNA:RNA Duplexes. *Nucleic Acids Res.* **1997**, *25*, 4429−4443.
(75) Brice, A. R.; Dominy, B. N. Analyzing the Robustness of the MM/PBSA Free Energy Calculation Method: Application to DNA Conformational Transitions. *J. Comput. Chem.* **2011**, *32*, 1431−1440.
(76) Hou, T. J.; Wang, J. M.; Li, Y. Y.; Wang, W. Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2011**, *51*, 69−82.
(77) Onufriev, A.; Case, D. A.; Bashford, D. Effective Born Radii in the Generalized Born Approximation: the Importance of Being Perfect. *J. Comput. Chem.* **2002**, *23*, 1297−1304.
(78) Owczarzy, R.; Moreira, B. G.; You, Y.; Behlke, M. A.; Walder, J. A. Predicting Stability of DNA Duplexes in Solutions Containing Magnesium and Monovalent Cations. *Biochemistry* **2008**, *47*, 5336−5353.
(79) Spink, C. H.; Garbett, N.; Chaires, J. B. Enthalpies of DNA Melting in the Presence of Osmolytes. *Biophys. Chem.* **2007**, *126*, 176−185.
(80) Nakano, S. I.; Karimata, H.; Ohmichi, T.; Kawakami, J.; Sugimoto, N. The Effect of Molecular Crowding with Nucleotide Length and Cosolute Structure on DNA Duplex Stability. *J. Am. Chem. Soc.* **2004**, *126*, 14330−14331.

DOI: 10.1021/acs.jpcb.5b09645
J. Phys. Chem. B 2015, 119, 15221−15234