Molecular modeling study of the effect of base methylation on internal interactions and motions in DNA and implication to B-Z conformation change

CAI Congzhong\textsuperscript{1,2}, WANG Wanlu\textsuperscript{1}, CHEN Yuzong\textsuperscript{1,2}

\textsuperscript{1}College of Mathematics and Sciences, Chongqing University, Chongqing 400044, P. R. China
\textsuperscript{2}Department of Computational Science, National University of Singapore, Singapore 117543

Received 18 November 2002; revised 24 February 2003

Abstract: Methylation in the bases of DNA is known to induce B-Z conformation change. In this work, molecular mechanics and normal mode analysis are used to probe how certain methylation affects the internal interactions and thermodynamic motions in the DNA double helixes in both B and Z conformations, and its implication to B-Z conformation change. By molecular modeling with Insight II, two cases involving cytosine C5 and guanine C8 methylation on both B and Z-form DNA duplex d(CGCGCG)\textsubscript{2} are studied in comparison with the corresponding unmethylated duplexes. The internal interaction energies computed based on a molecular mechanics force field and the entropies due to internal motions computed according to a normal mode analysis are in fare agreement with respective observed thermodynamic quantities. The analysis on the computed individual energy terms suggests that the observed B-Z conformation change induced by methylation is primarily driven by enthalpic factors. A combination of changes in Van der Waals interaction, electrostatic interaction and hydrogen bonding likely contributes to the change of enthalpy that favors Z-conformation in the methylated states.

Keywords: molecular modeling; DNA base methylation; conformation change; entropy; enthalpy

1. Introduction

Experimental studies have shown that a DNA molecule with alternating pyrimidine-purine sequence can adopt a left-handed, double-helical Z-DNA conformation \[1,2\]. Such structural changes of DNA occur as a consequence of environmental conditions such as at salt concentration above 4 mol/L NaCl \[1,2\] or through certain chemical modification such as methylation \[1\] or bromination \[2,3\] of bases at physiological conditions. These structural changes and chemical modifications are of importance in certain biological processes and thus are subjects of a number of investigations \[1-8\].

Molecular modeling has been separately used to construct methylated DNA structures \[4,5\], to analyze the contribution of internal interactions to the stability of DNA duplexes \[6,7\], and to examine the contribution of internal motions to B-Z transition \[8\]. The success of molecular modeling in these separate studies have raised an interesting question as to whether it can be used to probe the mechanism of the observed effect of methylation on DNA B-Z conformation change \[1,2\].

In the present work, in an attempt to probe possible mechanism of the effect of cytosine C5 and guanine C8 methylation on B-Z transition in DNA respectively, molecular modeling techniques are used to model the 3-dimensional structures of both normal and methylated DNAs. Effects of the methylation of every previously mentioned bases on the internal interactions and internal motions in both B and Z forms of DNA are analyzed to determine how they might affect the internal energy and other thermodynamic quantities which are important to determination of equilibrium state between different conformations. The results are used to assess the effect on B-Z conformation change.

2. Methods

2.1 Structures

The structures of both B- and Z-form d(CGCGCG)\textsubscript{2} were generated from Insight II \[9\] in which the standard parameters for helical rise and twist angle were used. The structures of both B and Z forms d(m\textsuperscript{5}CGCGm\textsuperscript{5}CG)\textsubscript{2} and d(CGm\textsuperscript{8}GCG)\textsubscript{2} were generated by adding methyl groups to the relevant positions of the corresponding B and Z form...
d(CGCGCG)\textsubscript{2} structures. Standard bond lengths and angles of Insight II were adopted to construct all those structures which were then optimized by 1000 iterations of steepest decent energy minimization using the Discover Module of Insight II.

2.2 Computation of internal interaction energies

The intra-molecular interaction energy of the DNA molecules was computed by the following empirical potential energy function that includes hydrogen bond, non-bonded Van der Waals and electrostatic interaction energy terms [10,11].

\[
V = \sum_{\text{H-bonds}} [V_0(l-\exp(-a(r-r_0)))^2 - V_0] + \sum_{\text{non-bonded}} \left[\frac{A_i}{r_{ij}^{12}} - \frac{B_i}{r_{ij}^6} + q_i q_j \varepsilon_{ij} r_{ij}\right] \tag{1}
\]

where \( r \) is the hydrogen bond donor-acceptor distance, \( V_0, a \) and \( r_0 \) are the hydrogen bond potential parameters; \( A_i \) and \( B_i \) are the nonbonded Van der Waals parameters; \( \varepsilon_i \) is the dielectric constant, \( q_i \) and \( q_j \) are the partial charges of the \( i \)-th and \( j \)-th atoms, and \( r_{ij} \) is the distance between them.

The nonbonded Van der Waals and electrostatic terms and their parameters were taken from the AMBER force field [12]. A distance-dependent dielectric constant [13] was used. In addition, to mimic the solvent effect on the phosphate-phosphate interactions at high salt concentrations, the dielectric constant between atoms of phosphate groups was taken as the geometric mean of the distance-dependent constant and that of water [14]. To avoid the difficulty in modeling the hydrogen dynamics, an implicit hydrogen atom Morse potential [15], which is a function of donor-acceptor distance, was used for hydrogen bond energy terms. This potential has been shown to give reasonable description of hydrogen bond energy and dynamics in bimolecular systems [10,11]. The published Morse potential parameters [11,16] were used in this work.

2.3 Normal mode computation

Under harmonic approximation, the internal motions of a biomolecule can be modeled by the following Hamiltonian:

\[
H_0 = \sum_{\text{atoms}} \frac{p_i^2}{2m} + \sum_{\text{bond-stretch}} \frac{1}{2} K_r (r-r_{eq})^2 + \sum_{\text{bond-angle}} \frac{1}{2} K_\theta \theta^2 + \sum_{\text{H-bond}} \frac{1}{2} K_{\sigma} (\theta-\theta_{eq})^2 + \sum_{\text{S-bond}} \frac{1}{2} K_{\phi} (\phi-\phi_{eq})^2 + \sum_{\text{non-bonded}} \frac{1}{2} K_{r_{ij}} (r_{ij}-r_{ij}^0)^2 + V_{\text{el}}
\]

where \( V_{\text{el}} \) is the static part of the Hamiltonian, i.e., the potential at an equilibrium position; \( K_r \) and \( r_{eq} \) are respectively the covalent bond (other than disulfide bond) force constant and equilibrium bond length; \( \theta \) is the bond angle; \( \theta_{eq} \) denotes the equilibrium bond angle; \( K_{\sigma}, K_{\phi}, K_{D}, K_{H} \) and \( K_{ij} \) are force constants for the bond angle bending, the bond rotating, the disulfide, the H-bond and the non-bonded interactions respectively; and \(<r>\) is the length in equilibrium positions. Since the changes in torsion and non-bonded forces are small, \( K_{\sigma} \) and \( K_{ij} \) can be given by the second derivative of the relevant potential; and they are from AMBER force field [12]. The force constants \( K_D \) and \( K_{H} \) can also be derived as the second derivative of the relevant potentials [10-12].

The equation of motions under the Hamiltonian in Eq.(2) was solved in mass weighted Cartesian coordinate system to give the vibrational frequency and eigenvector for each normal mode. The vibrational entropy due to internal motions was calculated by

\[
S_{v} = \left(\frac{R}{\pi}\right) \left[\sum_{l} \frac{\omega_l}{2RT} \coth\left(\frac{\omega_l}{2RT}\right) - \ln[2 \sinh\left(\frac{\omega_l}{2RT}\right)]\right] \tag{3}
\]

where \( R \) is the gas constant, the temperature \( T = 297 \text{ K} \), and \( \omega_l \) is the vibrational frequency for the \( l \)-th normal mode. The sum operator \( \sum_l \) is over all normal modes.

3. Results and discussion

In both B and Z forms, the overall structure of methylated DNA is found to be very similar to that of unmethylated DNA. Nonetheless, there is some structural variation in the backbone of the methylated DNA. Fig.1 shows the structure of B-form d(CGCGm8GCG)\textsubscript{2} superimposed onto that of B-form d(CGCGCG)\textsubscript{2}. Only local structural variation is found in the backbone region around the methylated G4 and G10 base in the methylated DNA, while the bases and the rest of the backbone is more or less unchanged. A similar behavior is found in the B-form d(m5CGCGm5CG)\textsubscript{2} where the structural distortion is relatively small and localized to the backbone of the methylated bases.
In the Z-form d(CG\textsubscript{m8}GCG\textsubscript{2}) and d(m\textsubscript{5}CGCGm\textsubscript{5}CG\textsubscript{2}), the structural variation is also relatively small and restricted to the backbone. However, the variation is not only restricted to the local backbone of the methylated base, but also extends to the backbone up to two bases away, which can be seen in Fig. 2 where the structure of Z-form d(m\textsubscript{5}CGCGm\textsubscript{5}CG\textsubscript{2}) super-imposed onto that of the Z-form d(CGCGCG\textsubscript{2}) is given. A similar phenomenon is also found in the Z-form d(CG\textsubscript{m8}GCG\textsubscript{2}).

Fig.1. Comparison between the modeled structure of B-DNA d(CGCGCG\textsubscript{2}) (in dashed lines) and that of the B-DNA d(CGm\textsubscript{8}GCG\textsubscript{2}) (in continuous lines). The methylated bases G\textsubscript{4} and G\textsubscript{10} in the second structure are marked by respective residue names.

As shown in Table 1, the measured entropy changes are (–53.6, –53.6, and –69.9) kJ/mol for d(CGCGCG\textsubscript{2}), d(m\textsubscript{5}CGCGm\textsubscript{5}CG\textsubscript{2}), and d(CG\textsubscript{m8}GCG\textsubscript{2}), respectively. The enthalpy change for the d(CG\textsubscript{m8}GCG\textsubscript{2}) DNA is approximately 16.3 kJ/mol higher than that for the other two DNAs. Its higher stability in the Z-form is thus enthalpically driven. From Table 1, the computed enthalpy changes are (–72.4, –80.4, and –104.2) kJ/mol for the respective structures, which are comparable with observed values. The computation seems to suggest that the larger enthalpy change in d(CG\textsubscript{m8}GCG\textsubscript{2}) results from a more reduced electrostatic repulsion among negatively charged phosphate groups.

Fig.2. Comparison between the modeled structure of Z-DNA d(CGCGCG\textsubscript{2}) (presented in dashed lines) and that of the Z-DNA d(m\textsubscript{5}CGCGm\textsubscript{5}CG\textsubscript{2}) (in continuous lines). The methylated bases C\textsubscript{1}, C\textsubscript{5}, C\textsubscript{7} and C\textsubscript{11} in the second structure are marked by respective residue names.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Form</th>
<th>Computed energy/(kJ/mol\textsuperscript{\dagger})</th>
<th>Computed energy difference/(kJ/mol\textsuperscript{\dagger})</th>
<th>measured</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(E) (E_V) (E_H) (E_e) (\Delta S_v) (\Delta S_H) (\Delta S_e) (\Delta E)</td>
<td>(\Delta E_V) (\Delta E_H) (\Delta E_e) (\Delta E_S) (\Delta E)</td>
<td></td>
</tr>
<tr>
<td>d(CGCGCG\textsubscript{2})</td>
<td>B</td>
<td>–771.5 –542.5 –285.0 56.5</td>
<td>–457.9</td>
<td>(-72.4) (-51.9) (-17.2) (-3.4)</td>
</tr>
<tr>
<td></td>
<td>Z</td>
<td>–843.9 –594.4 –302.2 53.1</td>
<td>–445.8</td>
<td>(-80.4) (-52.6) (-10.4) (-5.2)</td>
</tr>
<tr>
<td>d(m\textsubscript{5}CGCGm\textsubscript{5}CG\textsubscript{2})</td>
<td>B</td>
<td>–789.9 –561.8 –285.0 56.9</td>
<td>–474.7</td>
<td>(-80.4) (-52.6) (-10.4) (-5.2)</td>
</tr>
<tr>
<td></td>
<td>Z</td>
<td>–870.3 –644.6 –274.6 48.9</td>
<td>–491.4</td>
<td>(-104.2) (-78.7) (-30.1) (-69.9)</td>
</tr>
</tbody>
</table>

\textsuperscript{\dagger}E\_V, E\_H, E\_e and \(E\) represent respectively the Van der Waal, the H-bond and the electrostatic energies, and the sum of the first three. \(\Delta E\_V\), \(\Delta E\_H\), \(\Delta E\_e\) and \(\Delta E\) denote respectively the differences of according terms of B and Z-forms.

The atom positions of some of the phosphate groups in the Z-form methylated DNAs are markedly altered and thus induce changes in electrostatic as well as Van der Waals interactions. The effect of two closely located methylated bases on a strand in d(m\textsubscript{5}CGCGm\textsubscript{5}CG\textsubscript{2}) seems to interfere with and thus...
reduce the magnitude of the change in electrostatic interactions. As a result, its amplitude is substantially smaller than that of d(CGmGCGm8GCG)2 in which there is only one methylated base in each strand. Compared with unmethylated DNAs, the methylated Z-DNAs has a substantially more favorable Van der Waals interactions but slightly less favorable interbase hydrogen bond interactions. All these factors contribute to the overall enthalpy of the DNA.

The measured thermodynamic data suggests that d(CGCGCG)2 has an unfavorable entropy change towards the Z-form, while a favorable entropy change is expected for the d(CGmGCGm8GCG)2. An earlier normal mode study of B-Z transition in d(CGCGCG)2 gives a vibrational entropy change \( -T\Delta S \) of 27.6 kJ/mol [8]. The computed result from this study is 12.1 kJ/mol, which is comparable to that derived from the earlier study. Both computations indicate an unfavorable vibrational entropy change towards Z-conformation. In contrast, both the methylated DNAs has an approximately 16 kJ/mol favorable entropy change towards the Z-form.

References