Journal of Molecular Graphics and Modelling 29 (2011) 920–927

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

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

Effect of CpG methylation on DNA binding protein: Molecular dynamics simulations of the homeodomain PITX2 bound to the methylated DNA

Si-Ya Yang[a,*], Xiao-Li Yang[a], Li-Feng Yao[a], Hai-Bo Wang[b], Chen-Ke Sun[a]

[a] School of Chemistry and Chemical Engineering, Qujing Normal University, Sanjiang Road, Qinig District, Qujing City 655011, Yunnan Province, PR China
[b] School of Biological Resource and Environmental Science, Qujing Normal University, Qujing 655011, PR China

ARTICLE INFO

Article history:
Received 24 November 2010
Received in revised form 2 March 2011
Accepted 3 March 2011
Available online 10 March 2011

Keywords:
CpG methylation
DNA-binding PITX2
Molecular dynamics simulation
MM_GBSA
Binding free energy

ABSTRACT

A large number of studies have argued that aberrant CpG methylation is associated with some human cancers. One possible mechanism of the cancer caused by CpG methylation is the gene repression, which is a binding-inhibition of the sequence-specific transcription factors bound to specific DNA-binding sites. Exploring the effects of CpG methylation on the structure and the thermodynamic property of DNA-binding transcription factors will help to gain an insight into how CpG methylation affects the repression of gene transcription in cancer.

We have performed molecular dynamics (MD) simulations and free energy calculations of the protein bound to the native or the methylated DNA, in which the solution structure of the K50-class homeodomain PITX2 bound to DNA was used as a template. The simulation results indicate that the methylated CpG located at the DNA major groove can enhance the protein–DNA interactions, and the residue side-chains near the methylated CpG pair appear to have an unusually high affinity with DNA. The structural analysis and calculated energy values demonstrate that the binding-induced structural changes were further encouraged as the CpG methylation upon the complexation. Moreover, the CpG methylation may reduce the unfavorable effect of the conformational entropy and increase the electrostatic contribution to the binding free energy of DNA–PITX2. The changes in specific binding sites and the excessive affinity between DNA and protein caused by the CpG methylation could affect the gene transcriptional activity.

1. Introduction

DNA methylation is a crucial epigenetic modification in the genomes of higher eukaryotes. One of the most common forms of DNA methylation is 5-methyl-cytosine. In mammals, it is almost exclusively the CpG dinucleotide that can be methylated [1]. Bird suggested that proper DNA methylation is a prerequisite for normal cell development and can be involved in a variety of processes, such as gene repression, imprinting, X-chromosome inactivation and suppression of repetitive genomic elements [2]. Many studies have demonstrated that aberrant CpG methylation is associated with a large number of human diseases including cancer [3–8]. Furthermore, some recent studies show that DNA methylation is a potential biomarker for early detection and therapy monitoring of some cancers [9–11].

The role of DNA methylation in cancer development is associated with the repression of gene transcription. One possible mechanism for gene repression is the binding-inhibition of sequence-specific transcription factors attached to their specific DNA binding sites [12,13]. Another possible mechanism is the induction of a change in the state of the chromatin and the prevention of the transcriptional machinery from accessing the promoter region by proteins that recognize methylated DNA [13,14]. Reported examples include methyl-CpG binding proteins that specifically recognize methylated DNA. Petrovich and Veprintsev have studied the affinity difference for p53 bound directly to DNA sites containing singly or multiply methylated CpG dinucleotides. Their investigation revealed that the progressively increased methylation causes high affinity between DNA and protein. Other examples of DNA-binding proteins affected by DNA methylation include MLTF and CTCF, which do not directly bind methylated CpG [15]. However, no structural or energetic information is available on how CpG methylation might affect the interactions of DNA-binding protein. The present study focuses on characterizing the structural and thermodynamic impacts of methylated CpG on DNA–protein interactions. In this case, the protein does not directly contact to methylated CpG sites.

Homeodomains have been extensively studied [16–21] because they play a critical role in cellular processes and because they serve as a valuable model for probing the physical basis of protein–DNA

* Corresponding author. Tel.: +86 874 8998696; fax: +86 874 8998697.
E-mail addresses: yangsiyall@163.com (S.-Y. Yang), xlyang@yahoo.com.cn (X.-L. Yang), leoyao1982@yahoo.com.cn (L.-F. Yao), 13769862410@163.com (H.-B. Wang), Cksun99@126.com (C.-K. Sun).

1093-3263/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2011.03.003

interactions. Recently, Maiera et al. reported that the DNA methylation of the homeodomain transcription factor PITX2 has been validated as a biomarker for clinically relevant outcome prediction of early breast cancer patients [22,23]. The PITX2 can bind the special DNA core domain (TAATCC) adjacent to a CpG dinucleotide. Accordingly, The complex of PITX2 bound to a special DNA sequence could be made into a model that explores the impact of CpG methylation on protein–DNA interactions. In this paper, the solution structure of the K50-class homeodomain PITX2 bound to DNA (PDB ID: 1YZ8) was chosen as a template to construct a model which is the presumptive complex of PITX2-binding DNA containing methylated CpG. Based on the study of DNA-binding PITX2 by Rance and co-workers [20], we explored the interactions between PITX2 and DNA with methylated CpG using molecular dynamics simulation to reveal how CpG methylation might affect protein–DNA interactions.

Now molecular mechanics Poisson Boltzmann surface area (MM_PBSA) and molecular mechanics-generalized Born surface area (MM_GBSA) [24–26] protocols have become the most promising and widely used methods to estimate the binding free energy for protein–protein [27–30], protein–ligand [31–33], protein–DNA [34] and DNA–DNA systems [35]. In the MM_GBSA scheme, the binding free energy is estimated as the sum of the gas phase energies, solvation free energies and entropic contributions and is averaged over a series of snapshots from molecular dynamics (MD) trajectories. To obtain thermodynamic information about the PITX2 bound to the DNA and the energetic effects of CpG methylation on the DNA-binding protein, we estimated the binding free energies of the PITX2 with the native or the methylated DNA by using the MM_GBSA and GB^{OBC} [36] models.

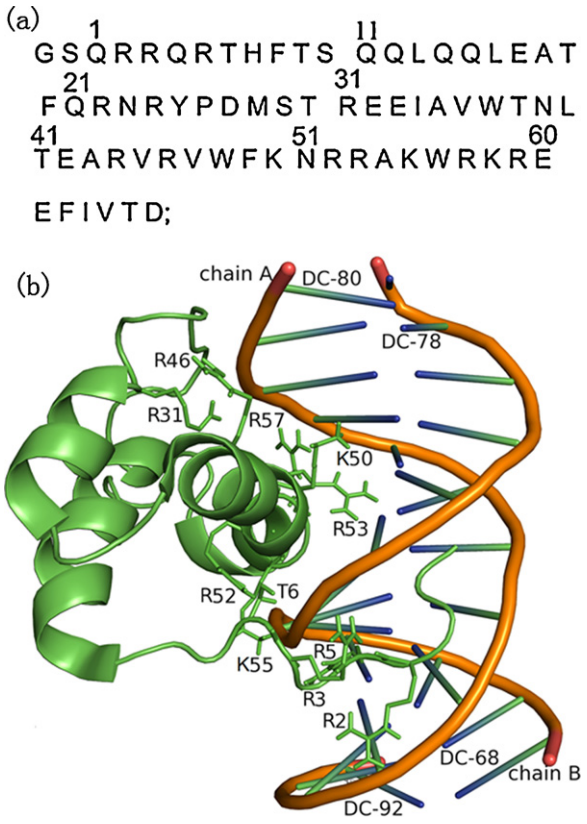

Fig. 1. Amino acid sequence of the PITX2 homeodomain used for the present studies (a). The cartoon diagram of PITX2 homeodomain–DNA complex (PDB ID: 1YZ8) (b). Showing the positions of the residues in the complex are some important hydrogen bonds of protein–DNA interaction in Ref. [20].

2. Method

2.1. Models with PITX2 bound to native and methylated DNA

In this study, the initial geometry for the PITX2–DNA complex, known as Pitxa, is one reported by Rance and co-workers. The DNA sequences used are as follows: 5′-G_{67} C_{68} T_{69} C_{70} T_{71} A_{72} A_{73} T_{74} C_{75} C_{76} C_{77} C_{78} G_{79}: C_{92} G_{91} A_{90} G_{89} A_{88} T_{87} T_{86} A_{85} G_{84} G_{83} G_{82} G_{81} C_{80}-3′. The amino acid sequence of the PITX2 homeodomain used is shown in Fig. 1(a) for the present studies. The analysis of the NMR structure of PITX2 bound to special DNA sites [20] is shown in Fig. 1(b). According to a study conducted by Iguchi-Arriga and Schaffner, in which the modified cytosine in the CpG motif is not directly in contact with the proteins [13], we constructed a complex of methylated DNA-binding protein PITX2, known as Pmc87, to carry out a MD simulation study by using the 1YZ8 structure as the template. In Pmc87, PITX2 is bound to the sequence-specific DNA containing a putative methylated step Me-$C_{78}G_{79}$/G_{81}Me-C_{80} (Me denotes a methyl group). To distinguish the effect of methylated 5′-CpG-3′ and 5′-GpC-3′ on the protein–DNA interaction, another presumptive complex containing a $G_{67}MeC_{68}$/MeC_{92} G_{91} step, known as Pmc96, was also constructed for a similar study.

The methylated cytosine was parameterized with the "general AMBER force field (GAFF03)" [37] and added to the residue library of AMBER. The atomic charges of 5-methylcytosine were computed using AM1-BCC method [38] in Antechamber tool.

2.2. Molecular dynamics simulation

We respectively performed six independent MD simulations for Pitxa, Pmc87 and Pmc96, using the AMBER program (Version 11) [39] and AMBER ff03 force field [40]. All molecules in the complexes were solvated by a cubic box of TIP3P water molecules [41] with a closeness parameter of 12 Å away from the boundary of any complex atoms. The each complex system was neutralized to neutrality adding nineteen Na^+ ions. After 2000 cycles of energy minimization, the systems were heated to 300 K and the density equilibration was performed at 300 K over 100 ps respectively. The weak restraints were maintained in each phase mentioned above. A constant pressure equilibration was then carried out for 1 ns at 300 K for each system. For each complex, we first performed a 5-ns MD simulations using the integration time step on 1 fs, respectively, and other five independent simulations for each complex were run up to 5-ns using the integration time step on 2 fs. The production phase of simulations was run for 5 ns using PMEMD (Particle Mesh Ewald Molecular Dynamics) [42] on the same condition as the final equilibration phase. Periodic boundary conditions were used with a constant number of particles, pressure, and temperature simulation criteria (NPT), and a cut-off distance of 10 Å. During the simulations, we wrote an output file every 500 steps and stored an actual frame every picosecond. All bonds involving hydrogen atoms were constrained by the SHAKE algorithm [43]. Subsequently, the dynamic behaviors and simulated structures for three complexes were analyzed by the root mean square deviation (RMSD).

2.3. Estimation of solvation energies and binding free energy

2.3.1. Physical model

According to Jelesarov et al. [34,44], prior to association, the free proteins and specific DNA sequence undergo a conformational change (structural adaptation, see Fig. 2) to the functional form. Accordingly, the thermodynamic cycle for the calculation of the binding free energy is presented in Fig. 3. All structural effects are accounted for explicitly by extensive MD sampling. At first, ensembles are generated in explicit water. Secondly, intermolecular and

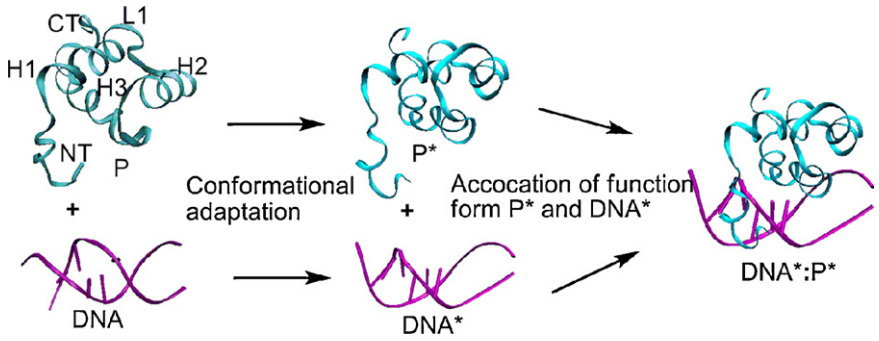

Fig. 2. Presumptive binding process for the interaction of PITX2 with the DNA. Free protein (P) and the B-DNA duplex (DNA) first undergo conformational transition to their functional form (DNA* and P*). In a second step, association involves formation (DNA*: P*) of intermolecular contacts across the preformed complementary binding interface. The final structures of the 5th ns MD trajectory are shown. Abbreviations: L1, loop 1; L2, loop 2; H1, helix 1; H2, helix 2; H3, helix 3; NT, N-terminus; CT, C-terminus.

intramolecular energies are computed using force field parameterization after removing the solvent, i.e., in the gas phase. Thirdly, the solvation free energy of each free species was calculated by GB^{OBC} (II) model and the total solvent accessible surface area (SA) [45] with appropriate parameters. The structures of unbound DNA and proteins in Fig. 2 are generated from the 1 ns trajectories of simulations that start from the structure extracted from the initial complexes and converge to the equilibrium state. All structures of free DNA (including native and methylation states) from MD trajectories are canonical B-form DNA, which may be thus used to represent the conformation of free DNA duplexes.

2.3.2. Free energy of conformational adaptation and of association of the structurally adapted proteins and DNA

In subcycle I of Fig. 3, the conformational adaptation (see Fig. 2) is depicted as a transition from a non-bound conformation (P and DNA) to the functional "binding-competent" form (P* and DNA*). The total conformational free energy in the aqueous phase ($\Delta G_{\mathrm{w}}^{\mathrm{conf}}$) is

$$\Delta G_{\mathrm{w}}^{\mathrm{conf}} = \Delta G_{\mathrm{MM}}^{\mathrm{conf}} + \Delta G_{\mathrm{sol,\,DNA^*+P^*}} - \Delta G_{\mathrm{sol,\,DNA+P}} \quad (1)$$

The molecular mechanics energy ($\Delta G_{\mathrm{MM}}^{\mathrm{conf}}$) is the summation of the differences ($\Delta \bar{E}_{\mathrm{ele}}^{\mathrm{conf}}$, $\Delta \bar{E}_{\mathrm{vdw}}^{\mathrm{conf}}$ and $\Delta \bar{E}_{\mathrm{int}}^{\mathrm{conf}}$) in each ensemble mean values of the electrostatic energy ($\bar{E}_{\mathrm{ele}}^{\mathrm{conf}}$), van der Waals energy ($\bar{E}_{\mathrm{vdw}}^{\mathrm{conf}}$), and internal energy ($\bar{E}_{\mathrm{int}}^{\mathrm{conf}}$) during the transition. In the last two terms of Eq. (1), $\Delta G_{\mathrm{sol,DNA+P}}$ is the solvation free energy for the free DNA and proteins, while $\Delta G_{\mathrm{sol,DNA^*+P^*}}$ is the corresponding value for the binding-competent conformations. The two terms are the sum of contributions from the polar solvation free energy ($\Delta G_{\mathrm{psol}}^{\mathrm{conf}}$)

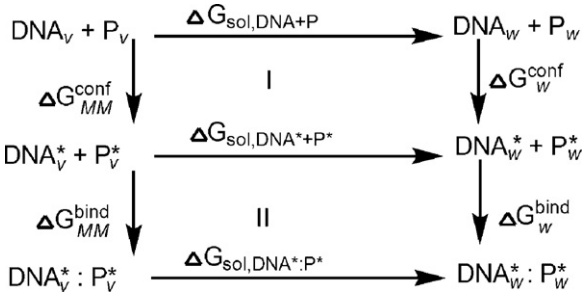

Fig. 3. Thermodynamic cycle used to calculate the free energy change of DNA binding PITX2. The subcycle I corresponds to structural transition from the initial state of the protein and DNA (P and DNA) to their binding-adaption conformations (P* and DNA*). The subcycle II represents the association of P* and DNA* to the DNA-protein complex. Left-hand and right-hand describe the binding process in the vacuum (with subscript *v*) and in the aqueous phase (with subscript *w*), respectively.

and the nonpolar solvation free energy ($\Delta G_{\mathrm{npsol}}^{\mathrm{conf}}$):

$$\Delta G_{\mathrm{sol,DNA^*+P^*}} - \Delta G_{\mathrm{sol,DNA+P}} = \Delta G_{\mathrm{psol}}^{\mathrm{conf}} + \Delta G_{\mathrm{npsol}}^{\mathrm{conf}} \quad (2)$$

The total electrostatic contribution ($\Delta G_{\mathrm{ele}}^{\mathrm{conf}}$) and the nonpolar contribution ($\Delta G_{\mathrm{np}}^{\mathrm{conf}}$) can be separately calculated as the sum of $\Delta G_{\mathrm{psol}}^{\mathrm{conf}}$ with $\Delta \bar{E}_{\mathrm{ele}}^{\mathrm{conf}}$ and the sum of $\Delta G_{\mathrm{npsol}}^{\mathrm{conf}}$ and $\Delta \bar{E}_{\mathrm{vdw}}^{\mathrm{conf}}$. Therefore, Eq. (1) can also be written as:

$$\Delta G_{\mathrm{w}}^{\mathrm{conf}} = \Delta G_{\mathrm{ele}}^{\mathrm{conf}} + \Delta G_{\mathrm{np}}^{\mathrm{conf}} + \Delta \bar{E}_{\mathrm{int}}^{\mathrm{conf}} \quad (3)$$

The energies relevant to Eqs. (2) and (3) were calculated using the AMBER molecular mechanical force field and GB^{OBC} (II) model in Amber 11 program. $\Delta G_{\mathrm{np}}^{\mathrm{conf}}$ is computed via $\Delta G_{\mathrm{surf}} = 0.005\,(\mathrm{kcal/mol}) \times A\,[\mathrm{\AA}^2]$, where A is the solvent-accessible surface of the molecule.

The association process of P* + DNA* → P*:DNA* is shown in Fig. 2 and subcycle II of Fig. 3. The associational free energy ($\Delta G_{\mathrm{w}}^{\mathrm{assoc}}$) is calculated by

$$\Delta G_{\mathrm{w}}^{\mathrm{assoc}} = \Delta G_{\mathrm{MM}}^{\mathrm{assoc}} + \Delta G_{\mathrm{sol,DNA^*:P^*}} - \Delta G_{\mathrm{sol,DNA^*+P^*}} \quad (4)$$

Nevertheless, the change in internal energy ($\Delta \bar{E}_{\mathrm{int}}^{\mathrm{assoc}}$) is zero for non-covalent association of P* + DNA* → P*:DNA*. Analogously to Eqs. (2) and (3), Eq. (4) can be written as:

$$\Delta G_{\mathrm{w}}^{\mathrm{assoc}} = \Delta G_{\mathrm{ele}}^{\mathrm{assoc}} + \Delta G_{\mathrm{np}}^{\mathrm{assoc}} \quad (5)$$

All calculations relevant to $\Delta G^{\mathrm{assoc}}$ were performed using the MM_GB module implemented in Amber Tool. Snapshots were taken at 10 ps time intervals from the corresponding MD trajectories. The explicit water molecules were removed from the snapshots.

2.3.3. Total binding free energy and entropic contribution

The total binding free energy is calculated as the sum of the conformational and associational free energies corrected by entropic effects:

$$\Delta G_{\mathrm{w}}^{\mathrm{bind}} = \Delta G_{\mathrm{w}}^{\mathrm{conf}} + \Delta G_{\mathrm{w}}^{\mathrm{assoc}} - T\Delta S_{\mathrm{tot}} \quad (6)$$

where ΔS_{tot} includes total entropic contributions accompanying complex formation and can be partitioned into:

$$\Delta S_{\mathrm{tot}} = \Delta S_{\mathrm{sol}} + \Delta S_{\mathrm{conf}} = \Delta S_{\mathrm{sol}} + (\Delta S_{\mathrm{vib}} + \Delta S_{\mathrm{tr}} + \Delta S_{\mathrm{rot}}) \quad (7)$$

In present study, ΔS_{sol}, as an entropic contribution related to solvent effects, is treated at different steps in the calculation of ΔG_{ele} and ΔG_{np}. ΔS_{conf} denotes the conformational entropy upon complexation [46], which is non-solvent-related contribution. In this paper, ΔS_{vib}, ΔS_{tr} and ΔS_{rot} are estimated by using NMODE module in Amber11 program MMPBSA.py. When one par-

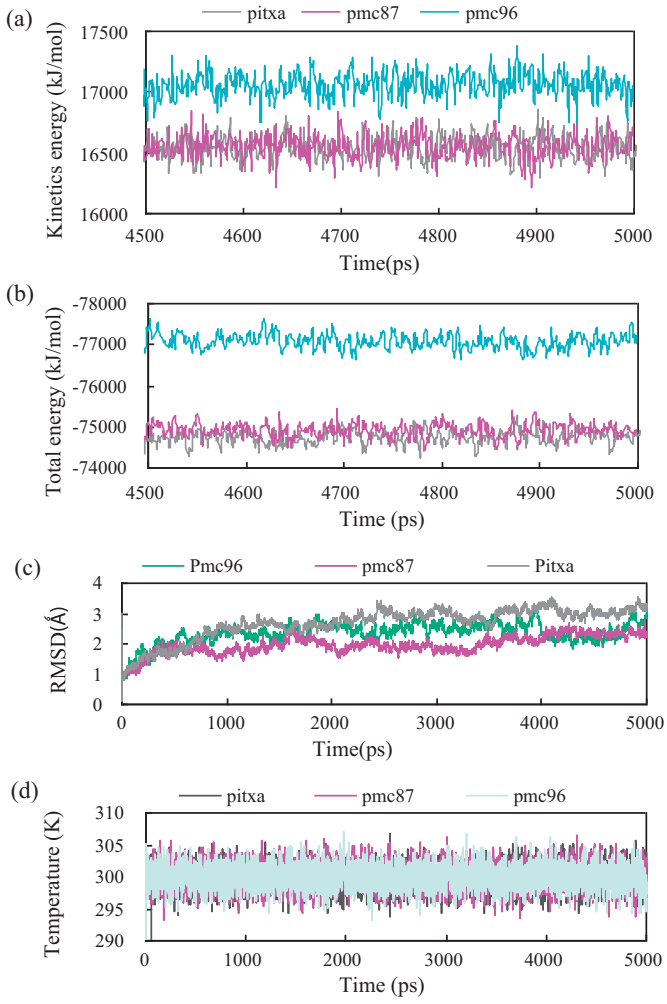

Fig. 4. The variation of kinetic energy of the complexes during the last 500 ps of simulation (a). The variation of total energy of the complexes during the last 500 ps of simulation (b). All atoms root mean square deviation (RMSD) of the complexes during the simulation (c). The variation of the temperature during 5 ns MD simulations (d). The integration time step is 1-fs through 5-ns MD simulations for each complex.

ticle (complex) is formed from two independent kinetic units, Eq. (6) can be thus expressed as:

$$\Delta G_{w}^{bind} = \Delta G_{w}^{conf} + \Delta G_{w}^{assoc} - T\Delta S_{conf} \quad (8)$$

3. Results and discussion

3.1. The simulated equilibrium of PITX2 bound to native and methylated DNA

As can be seen in Fig. 4(a), the kinetic energy components versus time in the final 500 ps of simulation is expected to fluctuate. The variation in total energy in the final 500 ps of the MD simulation shows small fluctuations through this time (see Fig. 4(b)). These facts indicate that the energy conservation is satisfied in the performed MD simulation [47]. After 2–3 ns of simulation, all RMSD values in three complexes changed slightly, as illustrated in Figs. 4(c) and S1. Obviously, the 5-ns simulation time employed is sufficient to obtain equilibrium structures of DNA–PITX2 complexes, and basic variation occurs in the latter geometry during last 2 ns. Therefore, the applied MD simulation specifies the geometries of the DNA–PITX2 complexes from the 5th ns MD trajectory in this paper. Moreover, the average temperature of 5 ns of simulation at 300 K for the studied systems was equal to 300 ± 1.78 K (Fig. 4(d)). The extracted equilibrium structures at 300 K belonging to the three complexes were obtained under temperature stable conditions.

3.2. Effects of the cytosine methylation on the PITX2-bound and the DNA-bound structures

The respective structural superimpositions of the P chain of 1YZ8 with those of PITXa, Pmc87 and Pmc96 exhibit a well-defined tertiary structure for PITX2 (see Supplementary information Fig. S2). When the six independent simulated structures of each complex are respectively superposed with 1YZ8, residues 3–58 have an averaged value of average root mean-square deviation (RMSD) from the 5th ns mean structure of 1.71 ± 0.20 Å for Pitxa, 1.70 ± 0.27 Å for Pmc87 and 1.70 ± 0.32 Å for Pmc96 for the backbone (N, Cα, C′, and O′). Obviously, there are no remarkable changes for conformations of PITX2 framework by the cytosine methylation. In the same process, the RMSDs from superposition structures of residues-2–6 (N-terminal i.e., NT) are 2.37 ± 0.38 Å for Pitxa, 2.35 ± 0.32 Å for Pmc87 and 1.81 ± 0.24 Å for Pmc96 respectively. The RMSDs from the superposition structure of residues 60–66 (C-terminal; i.e., CT) are 2.83 ± 0.43 Å for Pitxa, 2.06 ± 0.83 Å for Pmc87 and 2.32 ± 0.39 Å for Pmc96. The NT and CT segments have a sharp disorder in the three complexes, which is consistent with the reports of Rance and co-workers [20]. Interestingly, the NT segment of the PITX2 in Pmc96 and the CT segment of the PITX2 in Pmc87 are smaller deviate from those of 1YZ8 due to methylated cytosine.

The structural superposition of chain B or A of 1YZ8 with those of Pitxa, Pmc87 and Pmc96 were also analyzed in our study. For the backbone (P, O3′, O5′, C3′, C4′, C5′), the RMSDs of the structural superposition for chain B (nucleotides 67–79) are 2.00 ± 0.27 Å for Pitxa, 2.24 ± 0.49 Å for Pmc87, and 2.23 ± 0.71 Å for Pmc96. The backbone RMSDs for chain A (nucleotides 80–92) are 1.54 ± 0.17 Å for Pitxa, 1.65 ± 0.46 Å for Pmc87 and 1.54 ± 0.35 Å for Pmc96. The RMSDs of Chain B are much larger than that of Chain A, the cause of which could be stacking multiple pyrimidine-bases and weaker interactions between adjacent pyrimidine bases through the chain B. This implies that the distortion of the DNA conformation caused by cytosine methylation mostly occurs in the DNA segment with intensive pyrimidine-bases, which is in agreement with our conjecture from the employed *ab initio* calculation [48].

3.3. Effects of the cytosine methylation on the H-bond contacts between PITX2 and DNA

The structural analysis of the three simulated complexes indicates that, although there are some changes in the local tertiary structures for the PITX2-bound state by protein refolding, the relative position of three helixes for Pitx2 is generally consistent with the description by Rance and co-workers [20]. Therefore, our studies mainly focus on the impact of methylated cytosine on the interactions in binding region of DNA–PITX2, and the tertiary structure of the PITX2 homeodomain will not be discussed here.

All the simulated structures were run by the *ptraj* H-bond analysis using a distance cut-off of 3.5 Å and an angle cut-off of 120°. Some of crucial and stable H-bond contacts of DNA-binding protein PITX2 are listed in Table 1. How methylated cytosine might affect DNA–protein interactions are discussed below.

3.3.1. Effect of the cytosine methylation on PITX2 bound to the DNA minor groove

H-bond analysis, on six independent MD simulated structures for each complex, shows that there are stable hydrogen bond contacts between side chains of R2, R3, R5, Q4 and Ser-1 and DNA minor

Table 1
The significant H-bond contacts with the PITX2–DNA interface area of unmethylation and methylation complexes, the average distance (Å) and standard deviation (in parentheses).

Donor Res-atom	Acceptor Res-atom	Pitxa	Pmc96	Pmc87
G83-O6	Lys50-NZ	2.850(0.11)	2.873(0.14)	2.869(0.13)
G83-O6	Lys50-NZ	2.837(0.11)	2.844(0.13)	2.829(0.10)
G83-O6	Lys50-NZ	2.842(0.11)	2.860(0.14)	2.856(0.11)
G82-N7	Lys50-NZ	3.155(0.17)	2.873(0.14)[a]	3.065(0.16)
G82-N7	Lys50-NZ	3.087(0.16)	2.844(0.13)[a]	3.066(0.14)
G82-N7	Lys50-NZ	3.071(0.17)	3.151(0.18)[a]	3.036(0.14)
T74-O1P	Lys55-NZ	2.869(0.14)	3.014(0.29)	3.138(0.34)
C70-O2P	Lys58-NZ		2.883(0.16)	2.830(0.11)
T71-O1P	Arg59-NH1	2.824(0.11)	–	2.826(0.11)
T71-O1P	Arg59-NH2	2.834(0.12)	–	2.891(0.15)
G84-O2P	Arg57-NH1	2.909(0.15)	2.879(0.14)	2.909(0.15)
G84-O2P	Arg53-NH2	2.854(0.12)	3.102(0.17)	2.869(0.14)
A72-O2P	Arg52-NH1	2.807(0.10)	2.947(0.12)	2.807(0.10)
G81-O2P	Arg46-NH1	–	2.950(0.16)	3.081(0.22)
G81-O2P	Arg46-NH2	2.833(0.12)	–	2.948(0.22)
T74-O2P	Arg44-NH2	2.866(0.13)	2.832(0.12)	2.874(0.14)
G82-O2P	Arg31-NH2	2.854(0.12)	2.12(0.13)	2.837(0.13)
G82-O1P	Arg31-NH1	–	2.930(0.15)	2.903(0.15)
G82-O2P	Arg31-NE	–	–	2.878(0.14)
G89-O1P	Arg2-NH1	2.859(0.13)	2.869(0.13)	2.885(0.14)
G89-O2P	Arg2-NH1	2.885(0.14)	2.934(0.16)	2.934(0.16)
G89-O1P	Arg3-NE	2.934(0.16)	2.950(0.1)	2.944(0.15)
T71-O2	Arg3-NH1	2.918(0.14)	2.918(0.14)	2.901(0.14)
A72-N3	Arg3-NH1	3.352(0.04)	–	3.197(0.17)
G89-N3	Arg3-NH2	3.078(0.10)	–	3.153(0.16)
G89-O5′	Arg3-NH2	3.298(0.09)	2.833(0.12)[b]	3.130(0.18)
A73-O1P	Gln4-N	2.910(0.13)	2.839(0.10)	2.855(0.06)
T74-O1P	Arg5-NH1	2.894(0.15)	2.844(0.11)	2.894(0.15)
T74-O1P	Arg5-NH2	2.827(0.11)	2.800(0.10)	2.827(0.11)
T74-O1P	Arg5-NE	2.847(0.11)	2.894(0.15)	2.876(0.11)
T87-O2	Ser-1-OG	2.904(0.19)	2.743(0.15)	2.772(0.16)
T87-O2	Ser-1-N	2.910(0.13)	2.911(0.16)	2.930(0.13)

[a] H-bond distances between G84-N7 and Lys50-NZ.
[b] H-bond distance between A90-O5 and Arg3-NH2.

groove (see Table 1). Consistently, R2, R5 and Q4 appear to be making conserved contacts with phosphate backbone of the binding domain ($C_{70}T_{71}A_{72}A_{73}T_{74}$: $G_{89}A_{88}T_{87}T_{86}A_{85}$), while Ser-1 is making quit stable H-bonds with the atom O2 of base T87. These H-bond contacts are hardly affected by the CpG methylation.

Unlike residues mentioned above, H-bonds to R3 are also formed with atoms O2 and N3 of bases T71, A72 and G89 in Pitxa and Pmc87, respectively. When Arg3 lies approximately equidistant between the sugar–phosphate backbones of the two DNA strands in Pitxa and Pmc87, the guanidinium group of R3 facing the minor groove can be form multiple H-bonds with base atoms of the DNA CTAAT sequence. When the guanidinium group of R3 flanks to minor groove in Pmc96, H-bonds are formed with atoms O5′ of A90 (or atom O1P of A90) and atom O2 of T71. Based on the analysis above and the data in Table 1, we suggest that the H-bonds contacts of residue R3 with the minor groove play an important role for PITX2 recognizing the CTAAT sequence. The recognition is not influence by CpG methylation away from the minor groove. However, when the GpC is the methylation adjacent to the minor groove, the groove could be narrowed to cause the orientation change of the guanidinium of R3 so as to lose the H-bond contacts with base atoms at the minor groove bottom.

3.3.2. *Effect of the cytosine methylation on PITX2 bound to the DNA major groove*

In the simulated structures of three complexes, some H-bond contacts were found between the second and third helixes of PITX2 and DNA major groove (see Table 1).

In the second helix the H-bond contacts between the side chain of R31 and DNA were some different for three complexes. It is worth noting that, in Pmc87, the H-bond contacts between the side chain of R31 and DNA are more stable by the CpG methylation. In the third helix, R44, R52 and R53 appeared to be making conserved contact with the DNA backbone in three complexes. Compared with Pitxa and Pmc96, the more stable H-bonds were observed between the side chain of R46 and phosphate backbone groups of G82 for Pmc87 (see Table 1). Moreover, other H-bond contacts were respectively found between side chains of N51, K55, K58, R57 and R59 and backbone phosphate groups of DNA segment $C_{70}T_{71}A_{72}A_{73}T_{74}$ in difference complex.

Overall, the H-bond contacts between R44, R52 or R53 and the core sequence ($T_{71}A_{72}A_{73}T_{74}$: $A_{88}T_{87}T_{86}A_{85}$) are quite stable interactions upon PITX2–DNA complexations, which are hardly affected by the methylated cytosine outside the DNA–protein binding regions. It is important to note that, for Pmc87, the interactions of side chains in R46 and R31 with G82 represent an unusually high affinity by the CpG methylation. Previous studies have suggested that proteins with excessively high affinities for DNA or RNA can cause functional defects [49,50]. We may thus presume that the high affinity between R46 and R31 and G82 could affect some functions or properties of PITX2 transcription.

3.3.3. *Effect of the cytosine methylation on Lys50 (K50) recognize DNA*

H-bonds of side chain-base between lysine and guanine are very widespread in protein–DNA interaction [51,52]. Therefore, it is important to have a clear understanding of how cytosine methylation affects lysine–guanine interactions.

As can be seen from Table 1, the K50 side chain projects into the DNA major groove and forms multiple H-bonds with the atom N7 of G82 and the atom O6 of G83 in simulated structures of Pitxa and Pmc87. Both contacts between K50 protons and the N7 atom of G82 and the O6 atom of G83 have small fluctuations through the 5-ns MD simulation (see supplementary information Fig. S3). In Pmc96,

Table 2
The component of the energetic contributions to the free energy of PITX2 bound to the native and methylated DNA.[A]

Final complex	Pitxa	Pmc87	Pmc96
ΔG_{ele}^{conf}	1.18	−19.03	−7.51
ΔG_{np}^{conf}	−19.2	−3.24	13.2
$\Delta \bar{E}_{int}^{conf}$	87.38	118.07	74.58
ΔG_{w}^{conf}[B]	69.36	95.8	80.27
ΔG_{ele}^{assoc}	24.85	−13.42	7.90
ΔG_{np}^{assoc}	−159.52	−141.25	−143.88
ΔG_{w}^{assoc}[C]	−134.67 ± 1.05	−154.67 ± 1.14	−135.76 ± 4.89
$T\Delta S_{conf}$	−52.24 ± 2.01	−42.55 ± 1.43	−44.26 ± 2.30
ΔG^{bind}[D]	−13.07 ± 1.50	−16.33 ± 1.42	−11.23 ± 4.98

[A] All values are in kcal/mol.
[B] The $\Delta G_{w}^{conf} = \Delta G_{ele}^{conf} + \Delta G_{np}^{conf} + \Delta \bar{E}_{int}^{conf}$ (Eq. (3)).
[C] The $\Delta G_{w}^{assoc} = \Delta G_{ele}^{assoc} + \Delta G_{np}^{assoc}$ (Eq. (5)).
[D] Binding of free components (DNA and protein) to the final complex, binding free energy is $\Delta G^{bind} = \Delta G_{w}^{conf} + \Delta G_{w}^{assoc} - T\Delta S_{conf}$ (Eq. (8)). The definition of ΔG, ΔS and $\Delta \bar{E}$ with superscript and subscript in Table 2 see explanation in Section 2 of main text. For interrelated data see supplementary information Tables S1 and S2.

however, K50 would prefer to bind to the atom O6 of G83 and the atom N7 of G84, and both H-bond contacts of K50 NZ with G83 O6 and G84 N7 had obvious fluctuations throughout the 5-ns MD simulation.

Based on the results, we suggest that there is a high affinity between the K50 side chain and the G_{84} G_{83} G_{82} step of DNA. This affinity may be that the orientation of lysine is in an ideal position for the charged group to form H-bond contacts with the two guanines, and that atoms O6 and N7 of guanine are in the marked electronegative region of the major groove [53,54]. Therefore, the K50 side chain can form H-bond contacts with the $G_{84}G_{83}$ step [20] or the $G_{83}G_{82}$ step. This means that the side chain of K50 potentially mediates recognizing specific DNA sequence by linear fluctuating between multiple guanines and give priority to binding bases guanine adjacent to the methylated sites of DNA.

3.4. The binding free energy of PITX2 with the native and methylated DNA

The fact that cytosine methylation exerts a global effect on the interaction of the proteins bound to DNA is presented by several correlated alterations. To gain an insight into the mechanism involved, some energetic studies were carried out on the interactions of PITX2 with DNA containing methyl-cytosine. The correlative energies for three complexes were calculated based on the physical model sketched in Figs. 2 and 3. The calculated average values of electrostatic, van der Waals and internal energies ($\bar{E}_{ele}$, $\bar{E}_{vdW}$ and $\bar{E}_{int}$) and of electrostatic GB (G_{psol}) and non-polar solvation energies (G_{npsol}) are listed in Supplementary information Tables S1 and S2, respectively. The estimated results of DNA–PITX2 binding free energies for three complexes are reported in Table 2. Although the estimation of the free energies is approximate, the calculated results by the MM_GBSA and NMOD model appear very interesting on the energy difference of protein–DNA interactions caused by cytosine methylation. All calculations of MM_GBSA and NMOD were performed using the stable portion of the trajectories from 5th ns for six independent simulations of each complex.

3.4.1. The free energy of the DNA binding PITX2

The calculation results indicate that a conformational transition from DNA + P to DNA* + P* is a process of energy loss and entropy reduction for the Pitxa complex. The formation of DNA*:P* releases more associational free energy (ΔG^{assoc}), so that it can compensate the energetic penalty during the conformational adaptation. According to Eq. (8), the total binding free energy is −13.07 ± 1.50 kcal/mol, which approaches the calculated mean values for protein–DNA complexes by Jayaram et al. [44] and Gorfe and Jelesarov [34].

The energy values in Table 2 show that the total electrostatic interactions and the internal energies (ΔG_{ele} and $\Delta \bar{E}_{int}$) oppose the conformational transition to the Pitxa complexation, and the unfavorable contribution of electrostatic interactions essentially result from polar solvent energy (ΔG_{psol}). The non-polar interactions (ΔG_{np}), in which van der Waals interactions are predominant, drive DNA–PITX2 association. The conformational entropy is unfavorable for the association of PITX2–DNA, which is connected with the structural changes that PITX2 refolding caused in disorder and that the DNA produced the partial distortion for proteins bound to the DNA. The stabilizing role of the non-polar energy in Pitxa is in agreement with similar estimates for other protein–DNA complexes [55–59]. The energy values in Table 2 clearly reveal that PITX2–DNA binding induces structural changes upon complexation, and the changes deeply influence the distribution of the energetic component of the thermodynamic profile of DNA-binding protein PITX2.

3.4.2. Effect of the cytosine methylation on the free energy of the DNA-binding PITX2

We also studied the energetic effect of the cytosine methylation on DNA-binding PITX2. The associational energies (ΔG_{w}^{assoc}) for Pmc87 and Pmc96 were more favorable contributions to PITX2 bound to the methylated DNA (DNA* + P* → DNA*:P*). Compared with that of Pitxa, the pronounced change of ΔG_{w}^{assoc} mainly result from two contributions of favorable electrostatic and van der Waals interactions between PITX2 and methylated DNA.

During the conformational adaptation, the ΔG_{w}^{conf} of Pmc87 is 26.44 kcal/mol higher than that of Pitxa. This difference could result from the proteins refolding and the DNA distortion caused by bulk methyl groups on the CpG located at major groove. The ΔG_{w}^{conf} of Pmc96 is 10.92 kcal/mol higher than that of Pitxa but 15.53 kcal/mol lower than that of Pmc87. A possible reason is that the methylated GpC, adjacent to the minor groove and away from the DNA–PITX2 interface, induced only structural change within the N-tail arm and caused a slight effect on other residues of PITX2. The energy loss of adaptation process caused by the cytosine methylation is evidently a general trend, but the magnitude of energy consumed depended on methylated sites. The conformational entropies ($T\Delta S_{conf}$) of Pmc87 and Pmc96 are 9.69 kcal/mol and 7.98 kcal/mol positive than that of Pitxa respectively, which implies that the cytosine methylation can result in the change of conformational entropy so as to increase the favorable the contribution to PITX2–DNA association.

By analyzing data from Table 2, we suggest that the interaction of PITX2′s helix3 bound to the major groove is dominant during the complexation. When the CpG was methylated near the region of DNA-binding protein PITX2, the major groove could be widened to adapt to the bonds of protein side chains. In this case, the CpG methylation has scarcely any impact on the bonds of PITX2 inserted into the minor groove. The methylated GpC, by contrast, could inflict the minor groove narrowing to restrict the effective bonds of PITX2′s NT arm with the groove, but has no effect on the contacts of PITX2 with the major groove. These facts might be reason why the binding free energy increased by 3.26 kcal/mol for Pmc87 or only reduced by 1.84 kcal/mol for Pmc96, compared with that of Pitxa.

All analyses above indicate that the association of PITX2 with DNA is a typical reaction of binding-induced structural change. The associational reaction is mainly driven by non-polar interactions with a typical enthalpy behavior, but the conformation entropy counteracts binding. The CpG methylation, adjacent to the PITX2 binding domain (TAATC), reduces unfavorable influence of conformational entropy and enhances the contribution of electrostatic interactions to the association, The major reason could be that the

hydrophobic and steric effect of bulk methyl groups (i) induce the refolding of proteins to form a conformation adapting association, and (ii) cause the deformation of DNA helix to increase electrostatic interactions between phosphate backbone grounds and residues. During the association of PITX2 to the native DNA or the methylated DNA, the difference of dehydration of the non-polar surface is slight, but van der Waals packing interactions contribute a large amount of DNA-binding protein PITX2 when evaluated in a vacuum.

4. Conclusion

Based on calculated structures from six independent MD simulations of each complex, we carried out structural analyses of the PITX2–DNA interface and estimated the binding free energies. The results indicate that the crucial H-bond contacts between DNA and PITX2, as depicted in reference [20], are not greatly affected by the methylated cytosine, but the binding site and the stability of the H-bonds in different complexes differ to some degree. The methylated 5′-CpG, located at the DNA major groove, can causes the high affinity between DNA and PITX2 residues near the methylated sites and the enhancement of DNA–protein interaction. The 5′-GpC methylation near the DNA minor groove, in contrast, can results in the groove narrowing to restrict the effective bonds of PITX2′s NT arm with DNA minor groove. The binding free energy of the PITX2–DNA containing methylated cytosines primarily comes from the contribution of favorable electrostatic and hydrophobic interactions with enthalpy characteristics. The unfavorable entropy changes are compensated by the favorable enthalpy changes upon PITX2–DNA complexation.

In summary, the change of specific binding sites and excessive affinities of the DNA–protein association caused by the CpG methylation could be an important factor for the gene transcriptional inactivation.

Acknowledgements

This study is supported by National Natural Science Foundation of China (NNSFC) (No. 30730003) and Educational Natural Science Foundation of Yunnan Province of China (No. 07Z11574). We are grateful to Shanghai Supercomputer Center in China for the partial calculations of this paper and to Professor Tan Hong-Wei of Beijing Normal University for his guidance. The authors are particularly grateful to David A. Case Ph.D. in the Amber Development Team for his help.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2011.03.003.

References

[1] C.R. Calladine, H.R. Drew, B.F. Luisi, A.A. Travers, Understanding DNA: The Molecule and How it Works, Elsevier Academic Press, Amsterdam, 2004.
[2] A.P. Bird, DNA methylation patterns and epigenetic memory, Genes Dev. 16 (2002) 6–21.
[3] P.A. Jones, J.D. Buckley, The role of DNA methylation in cancer, Adv. Cancer Res. 54 (1990) 1–23.
[4] P.W. Laird, The role of DNA methylation in cancer genetics and epigenetics, Annu. Rev. Genet. 30 (1996) 441–464.
[5] R.L. Momparler, V. Bovenzi, DNA methylation and cancer, J. Cell. Physiol. 183 (2000) 145–154.
[6] M.C. Fruhwald, DNA methylation patterns in cancer: novel prognostic indicators? Am. J. Pharmacogenomics 3 (2003) 245–260.
[7] K.D. Robertson, DNA methylation and human disease, Nat. Rev. Genet. 6 (2005) 597–610, and references therein.
[8] C.A. Iacobuzio-Donahue, Epigenetic changes in cancer, Ann. Rev. Pathol. Mech. Dis. 4 (2009) 229–249.
[9] R.M. Brena, T.H.-M. Huang, C. Plass, Quantitative assessment of DNA methylation: potential applications for disease diagnosis, classification, and prognosis in clinical settings, J. Mol. Med. 84 (2006) 365–377.
[10] J. Tost, DNA methylation: an introduction to the biology and the disease-associated changes of a promising biomarker, Mol. Biotechnol. 44 (2010) 71–81.
[11] C. Kurkjian, S. Kummar, A.J. Murgo, DNA methylation: its role in cancer development and therapy, Curr. Prob. Cancer 32 (2008) 187–235.
[12] P.H. Tate, A.P. Bird, Effects of DNA methylation on DNA-binding proteins and gene expression, Curr. Opin. Genet. Dev. 3 (1993) 226–231.
[13] S.M.M. Iguchi-Arriga, W. Schaffner, CpG methylation of the cAMP-responsive enhance/promoter sequence TGACGTCA abolishes specific factor binding as well as transcripitional activation, Genes DEV. 3 (1989) 612–619.
[14] T. Vaissière, C. Sawan, Z.H. Epigenetic, interplay between histone modifications and DNA methylation in gene, Mutat. Res./Rev. Mutat. Res. 659 (2008) 40–48.
[15] M. Petrovich, D.B. Veprintsev, Effects of CpG methylation on recognition of DNA by the tumour suppressor p53, J. Mol. Biol. 386 (2009) 72–80.
[16] W.J. Gehring, M. Muller, M. Affolter, A. Percival-Smith, M. Billeter, Y. Qian, G. Otting, K. Wüthrich, The structure of the homeodomain and its functional implications, Trends Genet. 6 (1990) 323–329.
[17] C. Wolberger, Transcription factor structure and DNA binding, Curr. Opin. Struct. Biol. 3 (1993) 3–10.
[18] W.J. Gehring, M. Affolter, T. Burglin, Homeodomain proteins, Annu. Rev. Biochem. 63 (1994) 487–526.
[19] M. Priston, K. Kozlowski, D. Gill, K. Letwin, Y. Buys, A.V. Levin, M.A. Walter, E. Heon, Functional analyses of two newly identified PITX2 mutants reveal a novel molecular mechanism for Axenfeld-Rieger syndrome, Hum. Mol. Genet. 10 (2001) 1631–1638.
[20] A.C. Beth, C.B. Kimber, V. Dave, J. Ma, M. Rance, Solution structure of the K50 class homeodomain PITX2 bound to DNA and implications for mutations that cause Rieger syndrome, Biochemistry 44 (2005) 7497–7511.
[21] A.G. Toïth-Petroïczy, I. Simon, M. Fuxreiter, Y. Levy, Disordered tails of homeodomains facilitate DNA recognition by providing a trade-off between folding and specific binding, J. Am. Chem. Soc. 131 (42) (2009) 15084–15085.
[22] W.M. John, L.M. Astrid, S. Manfred, F. John, H. Nadia, DNA methylation as a biomarker in breast cancer, Future Oncol. 5 (2009) 1245–1256.
[23] S. Maiera, I. Nimmrich, T. Koeniga, et al., DNA-methylation of the homeodomain transcription factor PITX2 reliably predicts risk of distant disease recurrence in tamoxifen-treated, node-negative breast cancer patients – Technical and clinical validation in a multi-centre setting in collaboration with the European Organisation for Research and Treatment of Cancer (EORTC) PathoBiology group, Eur. J. Cancer 43 (2007) 1679–1686.
[24] J. Srinivasan, T.E. Cheatham, P. Cieplak, P.A. Kollman, D.A. Case, Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate–DNA helices, J. Am. Chem. Soc. 120 (1998) 9401–9409.
[25] P.A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D.A. Case, T. Cheatham, Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models, Acc. Chem. Res. 33 (2000) 889–897.
[26] D. Bashford, D.A. Case, Generalized Born models of macromolecular solvation effects, Annu. Rev. Phys. Chem. 51 (2000) 129–152.
[27] H. Gohlke, C. Kiel, D.A. Case, Insights into protein–protein binding by binding free energy calculation and free energy decomposition for the Ras–Raf and Ras–RalGDS complexes, J. Mol. Biol. 330 (2003) 891–913.
[28] H. Gohlke, L.A. Kuhn, D.A. Case, Changes in protein flexibility upon complex formation: analysis of Ras–Raf using molecular dynamics and a molecular framework approach, Proteins 56 (2004) 322–337.
[29] V. Zoete1, O. Michielin, Comparison between computational alanine scanning and per-residue binding free energy decomposition for protein–protein association using MM-GBSA: application to the TCR-p-MHC complex, Proteins: Struct. Funct. Bioinform. 67 (2007) 1026–1047.
[30] V. Zoetea, M.B. Irvingb, O. Michielin, MM–GBSA binding free energy decomposition and T cell receptor engineering, J. Mol. Recognit. 23 (2010) 142–152.
[31] J. Wang, P. Morin, W. Wang, P.A. Kollman, Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA, J. Am. Chem. Soc. 123 (2001) 5221–5230.
[32] I. Stoica, S.K. Sadiq, P.V. Coveney, Rapid and accurate prediction of binding free energies for Saquinavir-bound HIV-1 proteases, J. Am. Chem. Soc. 130 (2008) 2639–2648.
[33] G.Z. Ma, C.Q. Liu, Y.F. Qiu, J.M. Nan, Minor groove binding between DB818 and DNA: a molecular dynamics simulation and binding free energy analysis, Acta Chim. Sin. (chain) 67 (2009) 453–458.
[34] A.A. Gorfe, I. Jelesarov, Energetics of sequence-specific protein–DNA association: computational analysis of integrase Tn916 binding to its target DNA, Biochemistry 42 (2003) 11568–11576.
[35] S. Derreumaux, M. Chaoui, G. Tevanian, S. Fermandjian, Impact of CpG methylation on structure, dynamics and solvation of cAMP DNA responsive, Nucleic Acids Res. 29 (2001) 2314–2326.
[36] D. Qiu, P.S. Shenkin, F.P. Hollinger, W.C. Still, The GB/SA continuum model for solvation. A fast analytical method for the calculation of approximate born radii, J. Phys. Chem. A 101 (1997) 3005–3014.
[37] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollamn, D.A. Case, Development and testing of a general Amber force field, J. Comput. Chem. 25 (2004) 1157–1174.

[38] A. Jakalian, B.L. Bush, D.B. Jack, C.I. Bayly, Fast, efficient generation of high-quality atomic charges. AM1-BCC model I: method, J. Comput. Chem. 21 (2000) 132–146.
[39] D.A. Case, et al., AMBER 11, University of California, San Francisco, 2010.
[40] Y. Duan, C. Wu, S. Chowdhury, M.C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, P. Kollman, A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations, J. Comput. Chem. 24 (2003) 1999–2012.
[41] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935.
[42] T. Darden, D. York, L. Pedersen, Particle mesh Ewalds: an N log (N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[43] J.-P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Numerical integration of the Cartesian equations of motion in a system with constraints: Molecular dynamics of n-alkanes, J. Comput. Phys. 23 (1977) 327–341.
[44] B. Jayaram, K. Mconnell, S.B. Dixit, A. Das, D.L. Beveridge, Free-energy component analysis of 40 protein–DNA complexes: a consensus view on the thermodynamics of binding at the molecular level, J. Comput. Chem. 23 (2002) 1–14.
[45] D. Sitkoff, K.A. Sharp, B. Honig, Accurate calculation of hydration free energies using macroscopic solvent models, J. Phys. Chem. 98 (1994) 1978–1988.
[46] J. Numata, M. Wan, E.W. Knapp, Conformational entropy of biomolecules: beyond the quasi-harmonic approximation, Genome Inform. 18 (2007) 192–205.
[47] A.R. Leach, Molecular Modelling, Principles and Applications, second ed., 2001, pp. 390–392.
[48] Y.S. Chen, S.W. Yu, S.Y. Yang, Theoretic studies on the stacking interactions between DNA methylation–unmethylation bases, Acta Chem. (China) 8 (2010) 739–746.
[49] K. Watanabe, A.M. Lambowitz, High-affinity binding site for a group II intron-encoded reverse transcriptase/maturase within a stem-loop structure in the intron RNA, RNA 10 (2004) 1433–1443.
[50] M. Monsalve, B. Calles, M. Mencia, F. Rojo, M. Salas, Binding of phage [Phi] 29 protein p4 to the early A2c promoter: recruitment of a repressor by the RNA polymerase, J. Mol. Biol. 283 (1998) 559–569.
[51] Y. Mandel-Gutfreund, O. Schueler, H. Margalit, Comprehensive analysis of hydrogen bonds in regulatory protein–DNA complexes: in search of common principles, J. Mol. Biol. 253 (1995) 370–382.
[52] S. Perrier, J. Hau, D. Gasparutto, J. Cadet, A. Favier, J.-L. Ravanat, Characterization of lysine–guanine cross-links upon one-electron oxidation of a guanine-containing oligonucleotide in the presence of a trilysine peptide, J. Am. Chem. Soc. 128 (2006) 5703–5710.
[53] W. Saenger, Principles of Nucleic Acid Structure, Springer-Verlag, New York, 1984.
[54] P. Cysewski, P. Czeleń, Theoretical analysis of the effects of guanine oxidative damage on the properties of B-DNA telomere fragments, J. Mol. Model. 13 (2007) 739–750.
[55] M. Baginski, F. Fogolari, J.M. Briggs, Electrostatic and non-electrostatic contributions to the binding free energies of anthracycline antibiotics to DNA, J. Mol. Biol. 274 (1997) 253–267.
[56] V.K. Misra, J.L. Hecht, A.S. Yang, B. Honig, Electrostatic contributions to the binding free energy of the λcI Repressor to DNA, Biophys. J. 75 (1998) 2262–2273.
[57] B. Jayaram, K.J. McConnell, B.D. Surgit, D. Beveridge, Free energy analysis of protein–DNA binding: the EcoRI endonuclease–DNA complex, J. Comput. Phys. 151 (1999) 333–357.
[58] V.K. Misra, B. Honig, DNA binding properties of 2, 7-diazapyrene and its N-methylated cations studied by linear and circular dichroism spectroscopy and calorimetry, Proc. Natl. Acad. Sci. U.S.A. 92 (1995) 4691–4695.
[59] L. Jen-Jacobson, L.E. Engler, L.A. Jacobson, Structural and thermodynamic strategies for site-specific DNA binding protein, Structure 8 (2000) 1915–1923.