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.

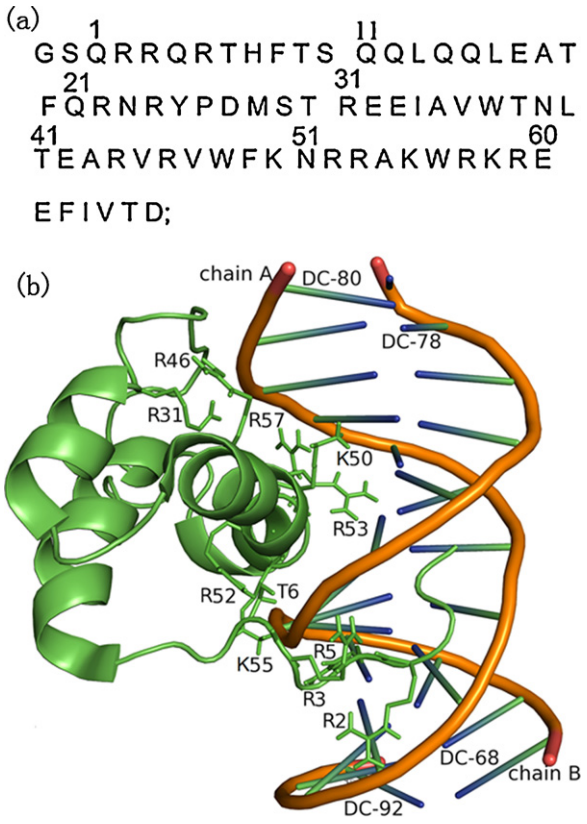

Fig. 1. Amino acid sequence of the PITX2 homeodomain used for the present studies (a). The cartoon diagram of PITX2 homeodomain–DNA complex (PDB ID: 1YZ8) (b). Showing the positions of the residues in the complex are some important hydrogen bonds of protein–DNA interaction in Ref. [20].

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