Electronic properties for detection of DNA methylation

Yigeng Tian, Zhongqi Liu, Yingying Cheng and Haiying Liu

School of Physics and Technology, University of Jinan, Jinan 250022, Shandong Province, P. R. China

1Email: ss_liuhy@ujn.edu.cn

Abstract. Detection of DNA methylation is of great significance for early diagnosis and precise treatments of related cancers. We calculated electronic properties of base pairs including cytosine (C), 5-methylcytosine (5mC), 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) with guanine (G). Calculation results show that modifications can regulate the energy gap of base pairs mainly by changing distributions of frontier orbitals and would not reduce the efficiency of charge transfer in DNA. 5fCG and 5caCG have better electrical transport properties due to the narrower gap and higher electron affinity compared with that of CG. Modified base pairs would remain stable when charge transport and not notably strengthen or weaken hydrogen bonding interactions of paired bases. Moreover, transverse electronic transport properties of benzamide with C and 5mC were studied by using the density functional theory (DFT) combined with the non-equilibrium green function (NEGF). The obtained current values of two systems have obvious difference under 0-0.2V biases, thus realizing the distinction of methylated cytosine.

1. Introduction

DNA methylation is a phenomenon that cytosine (C) is modified by methyl group in the 5th position of cytosine [1], which called 5-methylcytosine (5mC). In the process of demethylation, three intermediate products [2]: 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), 5-carboxylcytosine (5caC) are produced. DNA methylation belongs to epigenetic modification, which means it leads to gene silencing or overexpression without changing the type and quantity of bases, and its heredity has been reported in experiments [3]. It is well known that tumorigenesis is due to cell growth and differentiation disorders, which caused by gene mutation, deletion and amplification. However, many studies have found that DNA abnormal methylations are more common and important in the process of tumorigenesis [4], such as breast cancer, ovarian cancer, gastric cancer, colon cancer, bladder cancer and liver cancer [5,6]. Therefore, non-invasive detection of DNA methylation is helpful for early diagnosis and clarifying epigenetic mechanism of cancer [7].

Existing DNA methylation discrimination methods are mainly based on chemical DNA sequencing technologies. However, these methods generally have disadvantages of small screening area, high cost, complex design, low accuracy and difficult to control [8]. Besides, most of methods can not distinguish 5hmC from 5mC [9]. Many researches have found that 5hmC is involved in reprogramming of chromosome, transcriptional regulation of gene expression, and produces 5fC and 5caC, which are related to abnormal DNA methylation patterns in cancer [10-12] and expected to become biomarkers for early diagnosis of some cancers.
To explore a more convenient detection method of methylation, we studied differences in molecular structures and electronic properties of modified base pairs, thus indicating the effect of methylation and demethylation processes on DNA. Moreover, we calculated the transport properties of the related systems aimed to discriminate DNA methylation through measurable current values.

2. Computational methods
Base pairs of 5mC, 5hmC, 5fC, 5caC with guanine (G), their ionization potential and electron affinity, molecular junctions including benzamide [13] with C and 5mC were optimized at B3LYP/6-311++G** level. There is no symmetry restriction in optimization, and no imaginary frequency, which proves to be the lowest energy and stable structures. The binding energy of base pairs with single point energy correction were calculated by B3LYP method, and counterpoise correction (CP) scheme of Boys and Bernardi has been used to take into account of basis set superposition error (BSSE) [14]. All above calculations were completed by using Gaussian09 [15], and all calculated values are gas phase values. For the configuration of benzamide with C and 5mC, we have calculated transverse electronic transport properties using ATK (2015.01) [16], which is based on density functional theory (DFT) and non-equilibrium Green function (NEGF), the exchange-correlation function uses local density approximation (LDA.PZ) [17] and cutoff energy is 150 Ry. Nuclear electrons used Troullier-Martins nonlocal pseudopotential [18] and valence electrons used Siesta basis set, gold atoms used single-ζ plus polarization basis set (SZP) and other atoms used double-ζ plus polarization basis set (DZP) [19]. For the one-dimensional system with two electrodes that used periodic boundary conditions, the Brillouin zone adopted Monkhorst-pack method [20], and the K-point adopted (1×1×100). In the self-consistent calculation, the convergence value was set to 10⁻⁵ eV.

3. Results and discussion
3.1. Molecular configurations
We checked the bond length, bond angle and hydrogen bond length of optimized systems, and compared changes with the base pair of cytosine-guanine (CG). The number scheme for these base pairs are shown in Figure 1.

![Figure 1. Molecular configurations of CG, 5mCG, 5hmCG, 5fCG and 5caCG.](image)

Compared with C, the change ratios of bond lengths and bond angles of the four modified bases ranged from 0.00% to 1.32% and from 0.01% to 1.48%, and the major part of changes was focused on the modified C-5 site. Compared with CG pairs, hydrogen bond lengths of four modified base pairs range from 0.06% to 2.50%. It can be seen that these modifications have slight changes in the molecular
structures of bases, which is conducive to the stable existence of modified bases in DNA. In addition, the study on methylated-nonmethylated stacking bases by Chen showed that the degree of rotation and overlap between bases are increased by methylation of bases [21]. The π-π interaction between stacking bases and the hydrogen bonding are enhanced though methylation. This may be one of factors that lead to hypermethylation of cytosine in the CpG Island of promoter region of some cancer suppressor genes [22].

3.2. Electronic properties

3.2.1. Energy levels and frontier orbitals. We compared energy levels and frontier orbitals of CG, 5mCG, 5hmCG, 5fCG and 5caCG, that were shown in Figure 2 and Figure 3. It can be seen that the highest occupied molecular orbital (HOMO) of all base pairs are π orbitals delocalizated on guanine and basically unchanged, thus HOMO energy levels of each base pair are similar. In addition, the long-range charge transport in DNA is mainly achieved by hole hopping on the HOMO of CG [23], therefore existences of modified base pairs would not reduce the efficiency of charge transfer. Besides, their lowest unoccupied molecular orbital (LUMO) are π orbitals delocalizated on the side of cytosine, which have been changed due to the addition of modified groups, it can explain the changes of LUMO energy levels.

![Figure 2. Energy levels of CG, 5mCG, 5hmCG, 5fCG and 5caCG.](image)

![Figure 3. The HOMO-LUMO of CG, 5mCG, 5hmCG, 5fCG and 5caCG.](image)

Therefore, gap values of 5mCG and 5hmCG are slightly higher than CG, which is mainly caused by a slight increase of their LUMO energy levels. On the contrary, gap values of 5fCG and 5caCG are significantly decreased 0.77 eV and 0.64 eV, because of their HOMO energy levels only decrease 0.12
eV and 0.07 eV but their LUMO energy levels sharply decrease 0.89 eV and 0.71 eV, which are mainly caused by the distribution of their LUMO having been greatly changed and delocalization on modified groups. For 5fCG and 5caCG, the decrease of gap contributes to enhancement of their electrical conductivity.

### 3.2.2. Ionization potential, electron affinity, deformation energy and binding energy

We calculated the vertical ionization potential (VIP), adiabatic ionization potential (AIP), vertical electron affinity (VEA), adiabatic electron affinity (AEA), deformation energy of IP and EA (\(E_{\text{def}}\)), and binding energy (\(E_b\)) of all base pairs above, as shown in Table 1.

**Table 1.** The IP, EA, deformation energy, binding energy of CG, 5mCG, 5hmCG, 5fCG and 5caCG. All values are in eV.

| Base pairs | AIP  | VIP  | E_{\text{def}} | AEA  | VEA  | E_{\text{def}} | E_b  | E_{b}^{\text{BSSE}} |
|------------|------|------|---------------|------|------|---------------|------|-------------------|
| CG         | 6.92 (6.90[24]) | 7.29 | 0.36          | 0.41 | 0.02 | 0.39          | 1.05 | 1.00 (1.00[25])   |
| 5mCG       | 6.88 | 7.24 | 0.35          | 0.37 | 0.01 | 0.37          | 1.07 | 1.03              |
| 5hmCG      | 6.89 | 7.24 | 0.36          | 0.52 | 0.06 | 0.46          | 1.09 | 1.05              |
| 5fCG       | 7.03 | 7.41 | 0.37          | 1.22 | 0.75 | 0.48          | 0.94 | 0.90              |
| 5caCG      | 6.98 | 7.36 | 0.37          | 1.11 | 0.62 | 0.49          | 0.95 | 0.91              |

Calculation results show that the AIP and VIP of modified base pairs have a slight change with CG, that coincide with their HOMO energy levels. Moreover, AEA and VEA of 5fCG and 5caCG occur an obvious increase than CG, which means their ability of obtaining electrons is better than CG and are consistent with significant decreases of their LUMO energy levels. The magnitude of deformation energy reflects the relaxation degree of structure during obtaining or losing electrons, our calculated deformation energy of modified base pairs are similar to CG, thus these structures could remain stable when charge transport in DNA. Therefore, we predict that 5fCG and 5caCG should have great potential in molecular conductors. In addition, the definition of binding energy is \(E_b=E_{\text{base1}}+E_{\text{base2}}-E_{\text{base pair}}\), it represents strength of interaction between paired bases. According to calculated values, the stability of 5mCG and 5hmCG are minor better than CG, but 5fCG and 5caCG are limited weaker, it shows that modifications would not significantly strengthen or weaken hydrogen bonding interactions of paired bases.

### 3.3. Electronic transport properties

It was reported that benzamide can improve the recognition resolution of 5mC [13]. Therefore, we used benzamide as a probe and optimized the structure of benzamide connected with C and 5mC. Previous studies [26-28] have shown that there are minor differences in electrical properties whether bases connected with DNA backbone. In order to save computational effort, we removed the backbone.

![Figure 4. Schematic illustration of charge transport.](image-url)
Hydrogen atoms at para-position of acylamino group in benzamide were replaced by -SH groups, and that were inserted into fcc vacancy on the gold (111) electrodes surface. Dithiol was assumed to lose both hydrogen atoms when interacting with gold electrodes. The distance between sulfur atoms and gold interface was set at 0.19 nm, which is between reasonable range of 0.19-0.24 nm [29, 30], as shown in Figure 4. The transverse I-V curves are shown in Figure 5, it can be clearly seen that the current value of benzamide-5mC is as twice as the current value of benzamide-C under 0-0.2V bias, which can realize the distinction of C and 5mC by adding benzamide.

![Figure 5. I-V curves of molecular junctions for benzamide(B)-base(C and 5mC).](image)

4. Conclusions
Modifications existed in cytosine methylation and demethylation have slight influences on molecular structures of base pairs, which is conducive to their stability in DNA. Besides, modifications can regulate the energy gap of base pairs mainly by changing distributions of their LUMO, and the addition of modified groups should not significantly reduce the charge transfer properties of DNA. 5fCG and 5caCG have better electrical conductivities than CG due to their narrower gap and higher electron affinity. The deformation energy of modified base pairs are similar to CG, hence they could remain stable when charge transfer in DNA. In addition, changes of binding energy show that modifications would not notably strengthen or weaken hydrogen bonding interactions of paired bases. The current value of benzamide-5mC is as twice as the current value of benzamide-C under 0-0.2V biases, which can realize the detection of methylated cytosine. We think our research maybe provide meaningful theoretical support to recognition of DNA methylation.

Acknowledgments
This work was supported by the National Natural Science of Foundation of China (21673101) and National Undergraduate Innovation and Entrepreneurship Training Program (201710427004).

References
[1] Wu H and Zhang Y 2014 Reversing DNA Methylation: Mechanisms, Genomics, and Biological Functions Cell 156 45-68
[2] Kohli R M and Zhang Y 2013 TET enzymes, TDG and the dynamics of DNA demethylation Nature 502 472-9
[3] Jiang L, Zhang J, Wang J J, et al 2013 Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos Cell 153 773-84
[4] Ushijima T 2005 Detection and interpretation of altered methylation patterns in cancer cells Nat. Rev. Cancer 5 223-31
[5] Kulis M and Esteller M 2010 DNA methylation and cancer Adv. Genet 70 27-56
[6] Davis C D and Uthus E O 2004 DNA methylation, cancer susceptibility, and nutrient interactions Exp. Biol. Med. 229 988-95
[7] Sun K, Jiang P, Chan K C A, et al 2015 Plasma DNA tissue mapping by genome-wide methylation sequencing for noninvasive prenatal, cancer, and transplantation assessments Proc. Natl. Acad. Sci. 112 e5503-12
[8] Chao Z and Hailin W 2013 The progress on sequencing and detection of hydroxymethylated DNA Acta. Chim. Sin. 71 26-35
[9] Jin S G, Kadam S, Pfeifer G P 2010 Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine Nucleic. Acids. Res. 38 e125
[10] Chen Z X and Riggs A D 2011 DNA Methylation and Demethylation in Mammals J. Biol. Chem. 286 18347-53
[11] Smith Z D and Meissner A 2013 DNA methylation: roles in mammalian development Nat. Rev. Genet. 14 204-20
[12] Landan G, Cohen N M, Mukamel Z, et al 2012 Epigenetic polymorphism and the stochastic formation of differentially methylated regions in normal and cancerous tissues Nature Genet. 44 1207-14
[13] Huang S, He J, Chang S, et al 2010 Identifying Single Bases in a DNA Oligomer with Electron Tunnelling Nat. Nanotechnol. 5 868-73
[14] Boys S F and Bernardi F 1970 The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors Mol. Phys. 19 553-66
[15] Frisch M J, Trucks G W, Schlegel H B, et al 2009 Gaussian 09, Revision A. 02; Gaussian Inc: Wallingford, CT
[16] Brandbyge M, Mozos, José L, Ordejón, Pablo, et al 2002 Density-functional method for nonequilibrium electron transport Phys. Rev. B 65 165401
[17] Perdew J P and Zunger A 1981 Self-Interaction Correction to Density-Functional Approximations for Many-Body Systems Phys. Rev. B 23 5048-79
[18] Troullier N, Martins, José L 1991 Efficient pseudopotentials for plane-wave calculations Phys. Rev. B 43 1993-2006
[19] Soler J M, Artacho E, Gale J D, et al 2002 The SIESTA method for ab initio order-N materials simulation J. Phys.:Condens. Matter 14 2745
[20] Monkhorst H J 1976 Special points for Brillouin-zone integrations Phys. Rev. B 16 1748-9
[21] Chen Y, Yu S, Yang S 2010 Theoretical Studies on the Stacking Interactions between Methylated and Unmethylated DNA Bases Acta. Chim. Sin 68 739-46
[22] Widschwendter M, Jones P A 2002 The Potential Prognostic, Predictive, and Therapeutic Values of DNA Methylation in Cancer Clin. Cancer. Res 8 17-21
[23] Steinbrecher T, Koslowski T, Case D A 2008 Direct simulation of electron transfer reactions in dna radical cations J. Phys. Chem. B 112 16935-44
[24] Gervasio F L, Boero M, Parrinello M 2006 Double proton coupled charge transfer in DNA Angew. Chem. Int. Edit 45 5606-9
[25] Cheng Y, Liu H, Tian Y, et al 2019 Theoretical Study on Enhancement Effect of Amino Modification of Adenine on Conductivity of DNA Chem. J. Chinese University 40 279-87
[26] Staykov A, Tsuji Y, Yoshizawa K 2011 Conductance through Short DNA Molecules J. Phys. Chem. C 115 3481-90
[27] Liu H, Li G, Ai H, et al 2011 Electronic Enhancement Effect of Copper Modification of Base Pairs on the Conductivity of DNA J. Phys. Chem. C 115 22547-56
[28] Liu H, Li G, Zhang L, et al 2011 Electronic promotion effect of double proton transfer on conduction of DNA through improvement of transverse electronic communication of base pairs J. Chem. Phys 135 134315

[29] Liu H, Wang N, Zhao J, et al 2008 Length-Dependent Conductance of Molecular Wires and Contact Resistance in Metal-Molecule-Metal Junctions Chem Phys Chem 9 1416-24

[30] Cohen R, Stokbro K, Martin J M L, et al 2007 Charge transport in conjugated aromatic molecular junctions: Molecular conjugation and molecule–electrode Coupling J. Phys. Chem. C 111 14893-902