J Mol Model (2013) 19:1089–1098
DOI 10.1007/s00894-012-1628-4

ORIGINAL PAPER

Theoretical study on the binding mechanism between N6-methyladenine and natural DNA bases

Qi-Xia Song · Zhen-Dong Ding · Jian-Hua Liu · Yan Li · Hai-Jun Wang

Received: 19 June 2012 / Accepted: 3 October 2012 / Published online: 9 November 2012

Abstract N6-methyladenine (m^6A) is a rare base naturally occurring in DNA. It is different from the base adenine due to its N-CH_3. Therefore, the base not only pairs with thymine, but also with other DNA bases (cytosine, adenine and guanine). In this work, Møller-Plesset second-order (MP2) method has been used to investigate the binding mechanism between m^6A and natural DNA bases in gas phase and in aqueous solution. The results show that N-CH_3 changed the way of N6-methyladenine binding to natural DNA bases. The binding style significantly influences the stability of base pairs. The trans-m^6A:G and trans-m^6A:C conformers are the most stable among all the base pairs. The existence of solvent can remarkably reduce the stability of the base pairs, and the DNA bases prefer pairing with trans-m^6A to cis-m^6A. Besides, the properties of these hydrogen bonds have been analyzed by atom in molecules (AIM) theory, natural bond orbital (NBO) analysis and Wiberg bond indexes (WBI). In addition, pairing with m^6A decreases the binding energies compared to the normal Watson-Crick base pairs, it may explain the instability of the N6 site methylated DNA in theory.

Keywords Binding energy · DNA bases · Hydrogen bond · MP2 · N6-methyladenine

Introduction

Hydrogen bonding (H-bond) interaction plays a key role in biochemical systems, especially among nucleic acid base pairs [1]. The type of the H-bond has a direct influence on the secondary structure of DNA. Due to the importance, numerous researches focused on this aspect: the H-bond in normal base pairs [2, 3] and abnormal base pairs [4–6]. For example, Xue and Popelier investigated the substituent effects on Watson-Crick cytosine^{5X}:guanine, cytosine^{6X}:guanine [5] and guanine^{8X}:cytosine [6] base pairs.

There are kinds of abnormal bases which are called rare bases, most of them are methylated [7]. These methylated bases are usually in the form of C5-methylcytosine (m^5C), N6-methyladenine (m^6A) and N4-methylcytosine (m^4C) [8]. m^5C and m^6A are generally found in the genomes of many fungi, bacteria and protists, m^6A is also presented in archaeal DNA, whereas m^4C only exists in bacteria [9–11]. DNA methylation is relevant to gene expression, replication, repair [12] and genomic instability [13], and discovery of the key role of DNA methylation in regulation of genetic processes served as a principal basis and materialization of epigenetics and epigenomics [14, 15].

For all the methylated bases, researchers pay more attention to m^5C [16, 17], there has been less focused on m^6A. Actually, although m^6A exits in living beings as a minor base [18], it is considered to be the sixth base of DNA owing to the strong biological effects in bacteria while m^5C is the fifth [8, 9]. m^6A is connected with many fundamental biological processes such as cell differentiation and

Electronic supplementary material The online version of this article (doi:10.1007/s00894-012-1628-4) contains supplementary material, which is available to authorized users.

Q.-X. Song · Z.-D. Ding · J.-H. Liu · Y. Li · H.-J. Wang (✉)
The Key Laboratory of Food Colloids and Biotechnology, Ministry of Education, School of Chemical and Material Engineering, Jiangnan University, Wuxi, Jiangsu 214122, China
e-mail: wanghj@jiangnan.edu.cn

H.-J. Wang
e-mail: wanghj329@outlook.com

morphogenesis [19], and plays many important biological roles in DNA functions such as defense genetic signal in different bacteria [20], control bacterial virulence [21], but these mechanisms are still unclear [22, 23]. The m^6A base is also an important epigenetic signal for DNA replication and repair, protein-DNA interactions, host-pathogen interactions and other cellular processes [24].

In order to acquire more information about the functions of m^6A, the binding mechanism between m^6A and natural DNA bases has been systematic studied in this work. m^6A is different from the normal nucleotide base adenine due to its $N\text{-}CH_3$ part. Meanwhile, methylation of the amino group of adenine reduces the thermodynamic stability of DNA [25] and changes DNA curvature [26]. It indicates the methylation occurs at N-6 position of adenine, which would change the binding mechanism, namely, m^6A is not only paired with thymine (T), but also can mismatch with cytosine (C), adenine (A) and guanine (G).

Up to now, few reports involved in the binding mechanism between m^6A and natural DNA bases are delivered. In our work, a theoretical calculation has been performed to illustrate the effect for the methyl substitute of N6 site in gas phase and in aqueous solution. The H-bond characters and the binding energies of the mismatched base pairs have been studied, which may be conducive to study the structure of DNA, the pairing properties of the damaged base m^6A and the epigenetics.

Computational details

In this work, all the geometries of base pairs and free monomers were optimized using the second-order Møller–Plesset perturbational method (MP2) [27] with the 6-31G** basis set in vacuum and in aqueous solution. The solvent effect was considered using the polarized continuum model (PCM) of the self-consistent reaction field (SCRF) theory [28, 29]. No symmetry constraint was imposed during the optimization. Therefore, the geometry optimization for the saddle points occurred with all degrees of freedom. Each optimized structure was checked to be a true minimum through frequency calculations at the coincident level.

The binding energies of these base pairs were also calculated at the level of MP2/6-31G**. In addition, these binding energies (ΔE) were obtained by single point calculations using the individual optimized geometries as fragments in vacuum and in aqueous solution, respectively. The binding energy was evaluated as the difference between the total energy of a complex and the energies of its monomers [30]. To obtain more reliable binding energies, the binding energies are corrected by the basis set superposition error (BSSE) [31] via the counterpoise (CP) procedure method advanced by Boys et al. [32]. For systems under consideration, it can be calculated from Eq. 1:

$$\Delta E = E_{AB} - (E_A + E_B) + E_{BSSE}, \quad (1)$$

where E_{AB} is the single point energy of the base pair system; E_A and E_B are the single point energies of the m^6A and DNA base monomers, respectively; E_{BSSE} is the BSSE correction energy.

Subsequently the NBO [33] analysis and AIM theory analysis [34] were also performed at MP2/6-31G** in vacuum and in aqueous solution. All calculations were carried out within the framework of the Gaussian 03 program package [35], except that AIM analysis was employed by AIM2000 [36] package.

Results and discussion

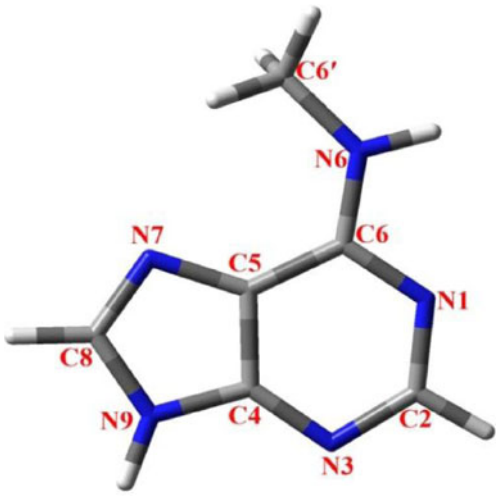
Geometries optimization of base monomers

For the sake of model simplification, the N9 and N1 hydrogen substituted bases were analyzed for purines and pyrimidines, respectively. The base monomers were originally optimized at MP2/6-31G** level from each initial guess in gas phase. We also consider the solvent effect, and the geometries were reoptimized in water solution using the PCM model. Vibrational frequency analysis on these optimized structures gave no imaginary frequencies suggesting that they were real minimum energy structures on the potential surfaces.

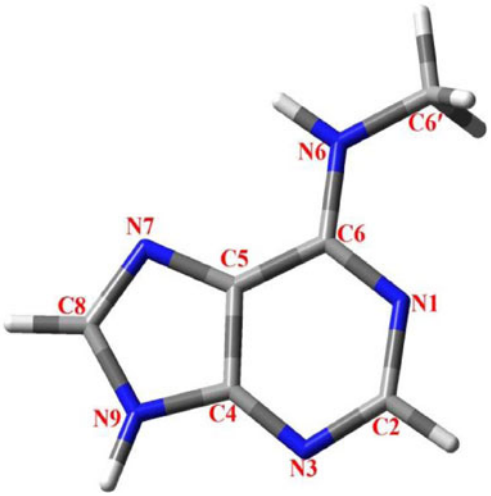
Bae and his co-workers reported that there were two forms of m^6A: trans-N6-methyladenine (trans-m^6A) and cis-N6-methyladenine (cis-m^6A) (Fig. 1), and the m^6A of a hemimethylated GATC site underwent a slow trans-cis interconversion [37]. Furthermore, both forms can detect in double helical DNA structure. The crystal structure of a single m^6A base showed that the methyl group attached at the adenine N6 position points toward the H-bond interface of the Watson-Crick base pair (the cis form), thereby, methylation of adenine N6 would alter the secondary structure of DNA [38]. However, structural studies of double-stranded oligonucleotides revealed that m^6A formed a normal Watson-Crick base pairing with the thymine in the complementary DNA strand (the trans form) [39]. All the investigations indicated that the dihedral angle $D_{N1\text{-}C6\text{-}N6\text{-}C6'}$ is closely related to the intermolecular H-bond interaction. To gain further insight into the structure of m^6A, potential energy surface scan of m^6A for the dihedral angle $D_{N1\text{-}C6\text{-}N6\text{-}C6'}$ increased with stepsize of 5.0°, and total 360.0° was carried out at the B3LYP/6-311++G** (Fig. 2) level. Two minima (-179.96° and 0.04°) are shown on the potential energy curves. Ultimately, two stable conformers (trans-m^6A and cis-m^6A) are obtained and

shown in Fig. 1. It can be concluded that when the dihedral angle $D_{N1\text{-}C6\text{-}N6\text{-}C6'}$ is close to planar, the stable geometry can be obtained. The difference value from the total energies of the two stable geometries is 12.44 kJ/mol, so both the stable conformers can coexist in the DNA structure. It is in good agreement with the experimental results [37]. Thus, both trans-m^{6}A and cis-m^{6}A should be taken into account for the initial geometries of the mismatched base pairs.

The optimized geometries and framework atom numbering of m^{6}A and natural DNA bases are shown in Fig. 1 and Fig. S1 (Supplementary information), respectively. The obtained geometries of trans-m^{6}A, cis-m^{6}A, A, T, G, and C are approximately planar.



trans- m^{6}A



cis- m^{6}A

Fig. 1 The optimized geometries for the bases calculated at the MP2/6-31G** level: trans-N6-methyladenine (trans-m^{6}A), cis-N6-methyadenine (cis-m^{6}A)

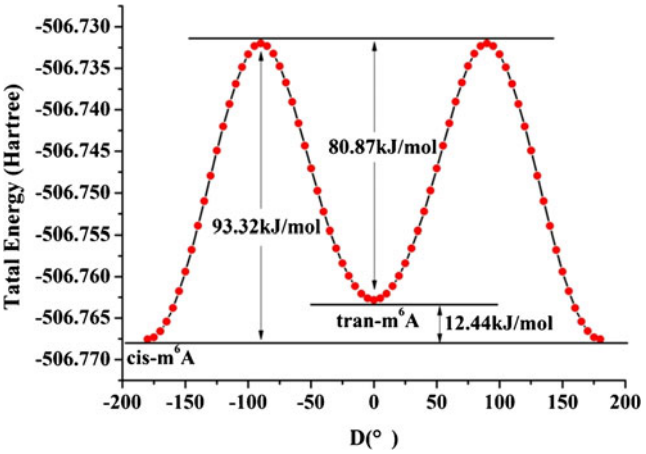


Fig. 2 Potential energy surface scan of m^{6}A for the dihedral angle $D_{N1\text{-}C6\text{-}N6\text{-}C6'}$: stepsize, 5.0°; total 360° was scanned (B3LYP/6-311++G**)

Geometries optimization, binding energies and NBO analysis of base pairs

The standard double helices are formed by two antiparallel strands (the glycosidic bonds are in cis orientation called Watson-Crick base pairs). However, DNA can also form parallel stranded (ps) double helix (the glycosidic bonds are in trans orientation called reverse Watson-Crick base pairs) [40–42]. To be consistent with the objective facts, all the geometries obtained are around the N1 site of m^{6}A. The representative optimized configurations are shown in Fig. 3 and Fig. S2, which were optimized in gas phase and in aqueous solution, respectively. The corresponding structural parameters of the base pairs are presented in Table 1, and the definitions of these parameters are presented in Fig. 4. As can be seen from the obtained results, all the mismatched base pairs' geometries make very great changes compared to the normal Watson-Crick base pairs [43], most of them become non-coplana, the trans-m^{6}A:A and cis-m^{6}A:A distorted mostly. In addition, the R became longer, $\alpha 1$ and $\alpha 2$ became smaller in aqueous solution.

The binding energies including BSSE correction of complexes calculated at MP2/6-31G** level in gas phase (E_G) and in aqueous solution (E_S) are given in Tables 2. As can be seen, the absolute values of the binding energies vary from 25.54 to 61.25 kJ/mol (in gas phase) and 16.74 to 36.11 kJ/mol (in aqueous solution). The results show that the binding energies of the pairs are largely reduced in aqueous solution, indicating that solvent existence can remarkably reduce the stability of the base pairs, and further destabilize the base pairs including the trans-m^{6}A.

The absolute values of the binding energies for the conformers trans-m^{6}A:G and trans-m^{6}A:C are 61.25 kJ/mol and 60.00 kJ/mol in gas phase, by comparison of the binding energies of other bases pairs, it showed that the trans-m^{6}A:G

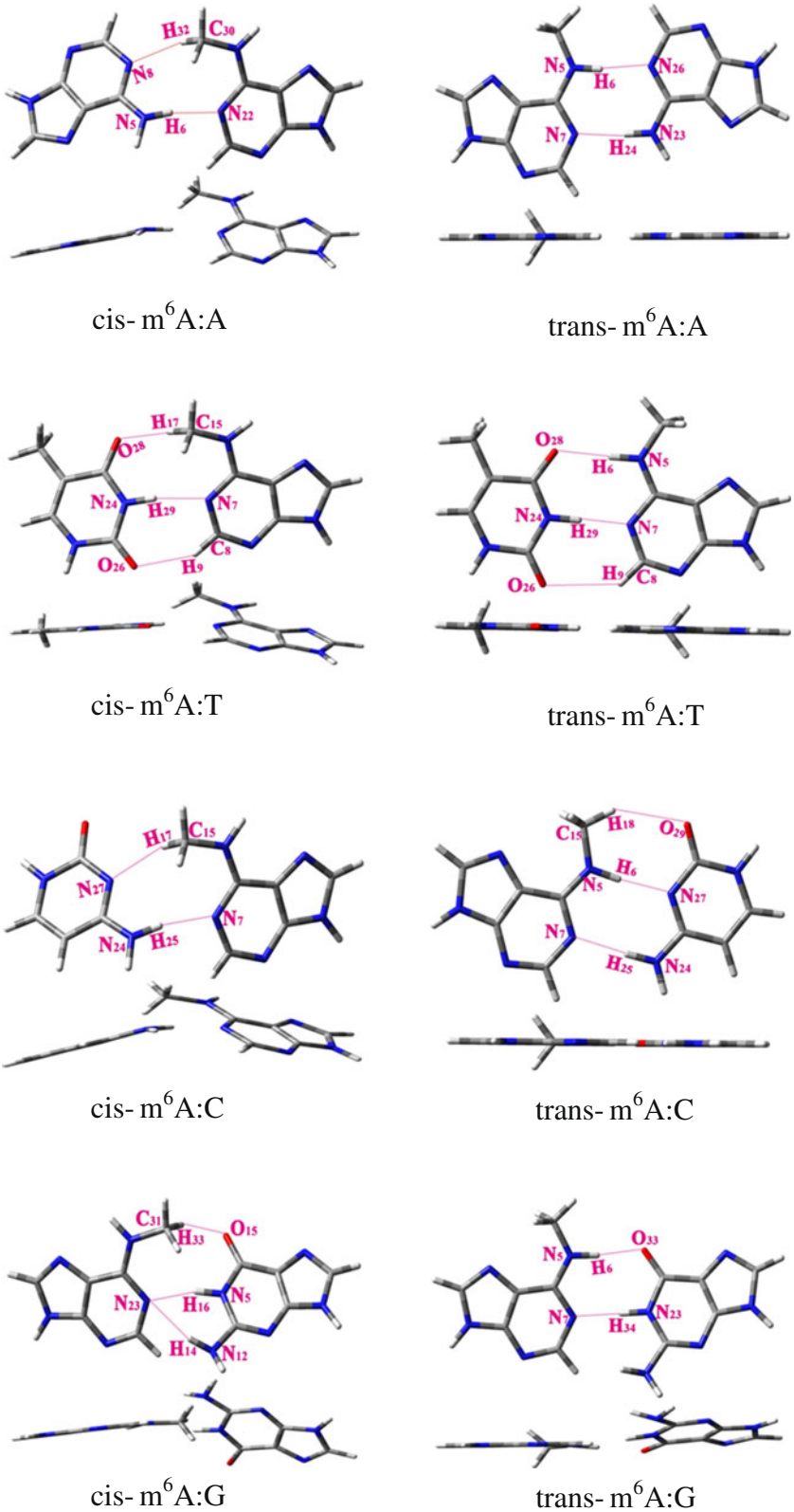


Fig. 3 The optimized configurations for base pairs including m^6A (trans and cis forms) calculated at the MP2/6-31G** level in gas phase. H-bonds are indicated by dotted line, and the corresponding atom numbering is given

and trans-m^6A:C are the most stable conformers among all the studied base pairs. In addition, the values are significantly lower than the binding energy of the normal Watson-Crick G:C base pair (the absolute values of the binding energy is 104 kJ/mol [44]). It can be concluded that when trans-m^6A paired with G and C, the binding ability of trans-m^6A can remarkably reduce the stability of the base pairs. The absolute values of the binding energies of the

Table 1 Structural parameters of the base pairs (R, in Å; α1, α2, in degrees)[a] calculated at MP2/6-31G** level in gas phase and aqueous solution

Base pair	Gas phase			Aqueous solution		
	R	α1	α2	R	α1	α2
cis-m^{6}A:A	12.58	52.23	29.23	12.51	53.45	29.41
cis-m^{6}A:T	9.52	63.60	64.04	9.62	61.51	60.97
cis-m^{6}A:C	11.15	45.29	37.22	11.40	44.58	34.15
cis-m^{6}A:G	11.96	54.16	50.47	12.01	52.60	49.31
trans-m^{6}A:A	12.97	30.41	27.51	12.98	30.05	27.16
trans-m^{6}A:T	9.95	56.67	56.94	10.01	55.48	55.84
trans-m^{6}A:C	11.28	19.70	38.24	11.30	18.93	37.79
trans-m^{6}A:G	12.20	46.11	47.50	12.22	45.91	46.44
A:T	10.00	54.94	56.35	10.06	53.94	55.41
C:G	10.17	53.17	55.52	10.16	52.62	55.35

[a] Definitions of R, α1 and α2 of the mismatched base pairs are presented in Fig. 4; For the definitions of R, α1 and α2 of the WC base pairs see ref [43]

conformers trans-m^{6}A:T and trans-m^{6}A:A are 50.68 and 46.56 kJ/mol, respectively. Meanwhile, the absolute value of the binding energy for the normal Watson-Crick A:T base pair is 57.16 kJ/mol [44]. Clearly, paired with T or A, the binding energies also reduced. A similar result was acquired in aqueous solution.

Obviously, both in gas phase and in aqueous solution, the absolute values of the binding energies of the base pairs shown in Table 2 are in an order of trans-m^{6}A:G ≈ trans-m^{6}A:C > trans-m^{6}A:T > trans-m^{6}A:A > cis-m^{6}A:G > cis-m^{6}A:T > cis-m^{6}A:C > cis-m^{6}A:A. It was observed that the absolute values of the binding energies of natrual DNA bases pairing with trans-m^{6}A are higher than pairing with the cis-m^{6}A. It can be concluded that the DNA bases are more favorable to pair with trans-m^{6}A than to pair with cis-m^{6}A. Thus, pairing with cis-m^{6}A further reduced the binding energy. To summarize, N6 methyl substituted adenine decreased the binding energies more or less both in trans and cis forms, it may be explained by the instability of the N6 site methylated DNA [25].

Table 2 Binding energies including BSSE correction of complexes (kJ/mol) calculated at MP2/6-31G** level in gas phase (E_G) and aqueous solution (E_S)

Base pair	E_G	E_S
cis-m^{6}A:A	−25.54	−16.74
cis-m^{6}A:T	−30.11	−21.24
cis-m^{6}A:C	−28.14	−17.17
cis-m^{6}A:G	−41.85	−26.77
trans-m^{6}A:A	−46.56	−31.00
trans-m^{6}A:T	−50.68	−34.78
trans-m^{6}A:C	−60.00	−36.11
trans-m^{6}A:G	−61.25	−36.02

To acquire more information about the H-bond interaction, the second-order interaction energies $E^{(2)}$ were obtained from NBO analysis in gas phase and in aqueous solution. The values of $E^{(2)}$ energies represent the different capacities of the donor-acceptor interaction for these base pairs to analyze the strength of H-bonds. The $E^{(2)}$ can be calculated from Eq. 2:

$$E^{(2)} = \Delta E_{ij} = q_i\left[F^2_{(i,j)}/\left(\varepsilon_i - \varepsilon_j\right)\right], \quad (2)$$

where q_i is the donor orbital occupancy, ε_i and ε_j are diagonal elements (orbital energies) and $F_{(i,j)}$ is the off-diagonal NBO Fock matrix element.

Returning to our investigated systems, the obtained $E^{(2)}$ energies of the proton donor and proton acceptor, along with the corresponding bond length and angles are listed in Table 3. It can be observed that the intermolecular binding type in all base pairs is formed between the electronegative N/O atoms and the active N-H or C-H groups. The N/O parts are the electron acceptors, and the N/C-H parts are the electron donors. Though C-H…N/O exist in non-Watson-Crick base pairs, the distance between the C-H and N/O is too far to be considered as a H-bond [45]. However, Koch believed C-H…N/O H-bonds exist in biomolecules which could be fruitful in understanding base pairing [46]. In our work, a series of theoretical methods were utilized to obtain reliable information and characterize C-H…N/O bonds.

H-bond will be indicated if the distance of the proton donor and proton acceptor is longer than the corresponding covalent bond distance and shorter than the sum of the van der Waals distance, and the corresponding angle is also greater than 90° [47]. It is clear that the bond distances of N-H…N/O and C-H…N/O vary from 1.79 to 2.49 Å and 2.26 to 2.69 Å in gas phase, 1.82 to 2.50 Å and 2.36 to 2.71 Å in aqueous solution, respectively; the corresponding angles are in the range of 130.57 to 179.59° and 135.87 to 168.73° in gas phase, 131.62 to 179.79° and 135.55 to 172.76° in aqueous solution, respectively, which implied that H-bonds have formed.

It also revealed that a stronger H-bond associated with a shorter length and a larger angle. Using the cis-m^{6}A:A in gas phase as reference, bond lengths, angles and $E^{(2)}$ energies of C_{30}-H_{32}…N_8 and N_5-H_6…N_{22} are 2.47 and 1.99 Å, 157.66 and 171.42°, 14.23 and 76.19 kJ/mol, respectively. It is obvious that the N_5-H_6…N_{22} has shorter length and larger angle than C_{30}-H_{32}…N_8, which implies that N_5-H_6…N_{22} is stronger than C_{30}-H_{32}…N_8. It can be concluded that strong H-bond prefers to be a straight angle and short bond length, meanwhile, N-H…N/O have shorter bond length and larger angles than that of C-H…O/N. Therefore, N-H…O/N are far stronger than C-H…O/N. Interestingly, in trans-m^{6}A:G, the $E^{(2)}$ energy of N_{12}-H_{14}…N_{23} is 5.48 kJ/mol, which seems to be opposite of the conclusion. In truth, the N_{12}-H_{14}…N_{23}

Table 3 The bond length (Å), angle (°) and corresponding second-order interaction energies $E^{(2)}$ (kJ/mol) for the base pairs calculated at MP2/6-31G** level in gas phase and aqueous solution

Base pair	Bond	Gas phase			Aqueous solution		
		Length	Angle	$E^{(2)}$	Length	Angle	$E^{(2)}$
cis-m^{6}A:A	C_{30}-H_{32}…N_8	2.47	157.66	14.23	2.60	164.45	9.41
	N_5-H_6…N_{22}	1.99	171.42	76.19	1.98	171.01	77.53
cis-m^{6}A:T	N_{24}-H_{29}…N_7	1.97	168.17	82.63	1.93	168.50	91.34
	C_{15}-H_{17}…O_{28}	2.26	168.73	20.13	2.36	172.76	13.60
	C_8-H_9…O_{26}	2.36	143.16	10.33	2.50	135.85	5.15
cis-m^{6}A:C	C_{15}-H_{17}…N_{27}	2.40	154.06	17.20	2.54	163.56	10.84
	N_{24}-H_{25}…N_7	1.99	168.22	75.02	1.98	170.52	78.78
cis-m^{6}A:G	N_5-H_{16}…N_{23}	1.97	164.18	76.36	1.90	166.23	101.50
	C_{31}-H_{33}…O_{15}	2.38	157.57	11.38	2.54	160.06	5.69
	N_{12}-H_{14}…N_{23}	2.49	130.57	5.48	2.50	131.62	4.31
trans-m^{6}A:A	N_5-H_6…N_{26}	1.97	174.63	85.52	1.98	175.69	84.14
	N_{23}-H_{24}…N_7	1.96	178.09	87.70	1.97	177.49	85.77
trans-m^{6}A:T	N_5-H_6…O_{28}	1.95	178.77	70.37	1.95	177.94	71.80
	N_{24}-H_{29}…N_7	1.79	179.59	156.52	1.82	179.79	140.46
	C_8-H_9…O_{26}	2.69	135.87	4.35	2.71	135.55	4.06
trans-m^{6}A:C	N_5-H_6…N_{27}	1.89	176.53	104.77	1.90	176.00	106.44
	N_{24}-H_{25}…N_7	1.95	177.87	93.60	1.99	176.93	81.00
	C_{15}-H_{18}…O_{29}	2.63	144.33	6.07	2.62	144.51	6.78
trans-m^{6}A:G	N_5-H_6…O_{33}	1.87	175.23	90.88	1.91	176.68	80.04
	N_{23}-H_{34}…N_7	1.85	177.43	131.80	1.85	179.08	132.72

bond length (2.49 Å) is too long and the angle (120.57°) is too small, thus weakening its strength.

Furthermore, the $E^{(2)}$ values of the N-H donor are relatively larger than that of C-H donor, it can be concluded that N-H…O/N are much stronger than C-H…O/N. Take the base pair cis-m^{6}A:A as an example, the $E^{(2)}$ values of the N_5-H_6…N_{22} and C_{30}-H_{32}…N_8 are 76.19 and 14.23 kJ/mol, respectively, which indicated that the strength of the donor-acceptor interaction of N_5-H_6…N_{22} is stronger than that of C_{30}-H_{32}…N_8. In addition, among all the base pairs, the $E^{(2)}$ value for N_{24}-H_{29}…N_7 in trans-m^{6}A:T is relatively larger than all the others, which exhibits the strongest H-bond capacity. Therefore, it can be concluded that N_5-H_6…N_{22} in cis-m^{6}A:A and N_{24}-H_{29}…N_7 in trans-m^{6}A:T contribute to their stabilities, respectively. The result is in good agreement with that discussed above.

From Tables 2 and 3 and Fig. 3, it can be easily observed that DNA bases coupled with trans-m^{6}A which involved two strong N-H…N/O H-bonds, however, there is only one N-H…N/O H-bond in the pairs which contain cis-m^{6}A. According to NBO analysis, the H-bond interaction of N-H…N/O bond is stronger than that of C-H…N/O, which has greater $E^{(2)}$. Take the trans-m^{6}A:A and cis-m^{6}A:A for example, the absolute values of the binding energies are 46.56 and 25.54 kJ/mol, respectively. The pair trans-m^{6}A:A contains N_5-H_6…N_{26} ($E^{(2)}$, 85.52 kJ/mol) and N_{23}-H_{24}…N_7 ($E^{(2)}$, 87.70 kJ/mol) while the pair cis-m^{6}A:A contains N_5-H_6…N_{22} ($E^{(2)}$, 76.19 kJ/mol) and C_{30}-H_{32}…N_8 ($E^{(2)}$, 14.23 kJ/mol). It is also found that N-H…N/O in the base pairs with trans-m^{6}A is stronger than that with cis-m^{6}A, revealing that the activity of N-H is weakened by the methyl, and became much lower in cis-m^{6}A. Therefore, the trans form participated in conformers is more stable than cis form which offers a remarkable insight into explaining the reason.

Moving to solvated systems, the solution effect results in a longer H-bond length and a smaller angle. The values of $E^{(2)}$ in trans pairs reduce a lot, which indicates that the H-bonds become weaker in aqueous solution. Notwithstanding, the values of $E^{(2)}$ in cis pairs, the N-H…N/O becomes stronger while the C-H…N/O becomes weaker, meanwhile, the binding energies are largely reduced in aqueous solution, which suggests that C-H…N/O plays a key role in the base pairs.

AIM and NBO analysis of base pairs

In order to obtain more information about the H-bonds, the AIM theory was used to analyze the bonding characteristics at the MP2/6-31G** level in gas phase and in aqueous solution. AIM theory provides a universally applicable tool for the classification of the bonding interactions that take place in any molecular system, even inside a supermolecule [48], which is based on the topological analysis of the properties of the electron density (ρ_c) and its Laplacian of electron density ($\nabla^2\rho_c$) at bond critical points (BCPs). The

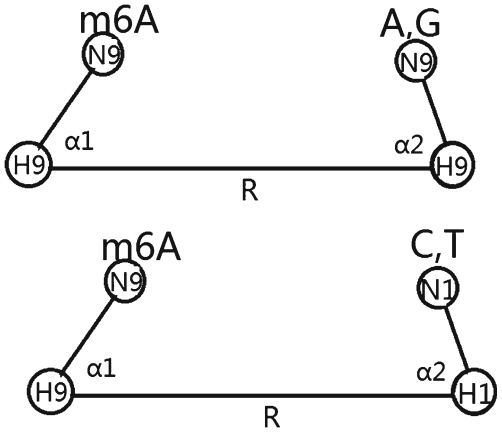


Fig. 4 Definitions for the geometrical parameters (R, α1, α2 between every two bases) of the mismatched base pairs

ρ_c value is used to describe the bond strength; a stronger bond is associated with a larger ρ_c value. The $\nabla^2\rho_c$ value describes the characteristic of the bond. If $\nabla^2\rho_c < 0$, it is named as the covalent bond; If $\nabla^2\rho_c > 0$, it refers to a closed-shell interaction and the characteristic of an ionic bond, hydrogen bond or van der Waals interaction. Small and positive values of $\nabla^2\rho_c$ indicate that a small charge concentration takes place along the bond path linking two nuclei, and a large electron density at the bond critical point and a positive value of $\nabla^2\rho_c$ indicate a strong H-bond [49]. There are a set of criteria for ρ_c and $\nabla^2\rho_c$ proposed at BCPs for the conventional H-bonds. Both parameters for closed-shell interactions as H-bonds are positive within the following ranges: 0.002-0.040 a.u. for the electron density and 0.024-0.139 a.u. for its Laplacian [50]. The AIM analysis of the base pairs with BCPs are shown in Fig. 4 (in gas phase) and Fig. S3 (in aqueous solution), the corresponding ρ_c and $\nabla^2\rho_c$ values for the H-bonds are listed in Table 4. For most H-bonds considered here, the ρ_c and $\nabla^2\rho_c$ values lie in the relative proposed ranges. It can be observed that ρ_c and $\nabla^2\rho_c$ at BCPs of H-bonds fall within 0.0060-0.0412 a.u. and 0.0234-0.1122 a.u. in gas phase, 0.0058-0.0382 a.u. and 0.0224-0.1062 a.u. in aqueous solution respectively. It can be concluded that the interactions are all closed shell systems (H-bond interaction) Fig. 5.

According to Tables 2, 3 and 4, it can be found that with greater binding energy, the conformer has bigger ρ_c and $\nabla^2\rho_c$ values. For example, as the most stable base pairs, the biggest binding energies belong to trans-m^{6}A:G and trans-m^{6}A:C, and there are two strong H-bonds N_5-H_6…O_{33} (ρ_c, 0.0294 a.u. and $\nabla^2\rho_c$, 0.0989 a.u.), N_{23}-H_{34}…N_7 (ρ_c, 0.0353 a.u. and $\nabla^2\rho_c$, 0.1034 a.u.); and N_5-H_6…N_{27} (ρ_c, 0.0324 a.u. and $\nabla^2\rho_c$, 0.0954 a.u.), N_{24}-H_{25}…N_7 (ρ_c, 0.0288 a.u. and $\nabla^2\rho_c$, 0.0798 a.u.) which belonged to the conformers, respectively. In addition, the behavior of $\nabla^2\rho_c$ is parallel to that exhibited by ρ_c. It is clear that the larger ρ_c and $\nabla^2\rho_c$ values contribute to stronger H-bonds, moreover, strong H-

Table 4 Wiberg bond indexes (WBI) and properties of the electron density of bond critical point (ρ and $\nabla^2\rho_c$, in a.u.) for the base pairs calculated at MP2/6-31G** level in the gas phase and aqueous solution

Base pair	Bond	Gas phase			Aqueous solution		
		WBI	ρ_c	$\nabla^2\rho_c$	WBI	ρ_c	$\nabla^2\rho_c$
cis-m^{6}A:A	C_{30}-H_{32}…N_8	0.0071	0.0115	0.0333	0.0055	0.0090	0.0264
	N_5-H_6…N_{22}	0.0351	0.0264	0.0760	0.0361	0.0268	0.0770
cis-m^{6}A:T	N_{24}-H_{29}…N_7	0.0434	0.0273	0.0766	0.0465	0.0294	0.0838
	C_{15}-H_{17}…O_{28}	0.0103	0.0145	0.0443	0.0089	0.0124	0.0377
	C_8-H_9…O_{26}	0.0065	0.0119	0.0398	0.0046	0.0093	0.0331
cis-m^{6}A:C	C_{15}-H_{17}…N_{27}	0.0081	0.0132	0.0385	0.0068	0.0103	0.0295
	N_{24}-H_{25}…N_7	0.0357	0.0261	0.0750	0.0366	0.0268	0.0773
cis-m^{6}A:G	N_5-H_{16}…N_{23}	0.0401	0.0278	0.0785	0.0512	0.0321	0.0908
	C_{31}-H_{33}…O_{15}	0.0087	0.0118	0.0370	0.0060	0.0088	0.0291
	N_{12}-H_{14}…N_{23}	0.0058	0.0103	0.0352	0.0054	0.0101	0.0345
trans-m^{6}A:A	N_5-H_6…N_{26}	0.0410	0.0277	0.0780	0.0404	0.0273	0.0763
	N_{23}-H_{24}…N_7	0.0424	0.0281	0.0793	0.0414	0.0278	0.0783
trans-m^{6}A:T	N_5-H_6…O_{28}	0.0338	0.0244	0.0785	0.0349	0.0246	0.0783
	N_{24}-H_{29}…N_7	0.0775	0.0412	0.1122	0.0692	0.0382	0.1062
	C_8-H_9…O_{26}	0.0029	0.0060	0.0234	0.0031	0.0058	0.0224
trans-m^{6}A:C	N_5-H_6…N_{27}	0.0484	0.0324	0.0954	0.0511	0.0321	0.0922
	N_{24}-H_{25}…N_7	0.0469	0.0288	0.0798	0.0397	0.0265	0.0737
	C_{15}-H_{18}…O_{29}	0.0039	0.0069	0.0255	0.0047	0.0074	0.0264
trans-m^{6}A:G	N_5-H_6…O_{33}	0.0454	0.0292	0.0965	0.0403	0.0270	0.0770
	N_{23}-H_{34}…N_7	0.0657	0.0363	0.1002	0.0669	0.0364	0.0838

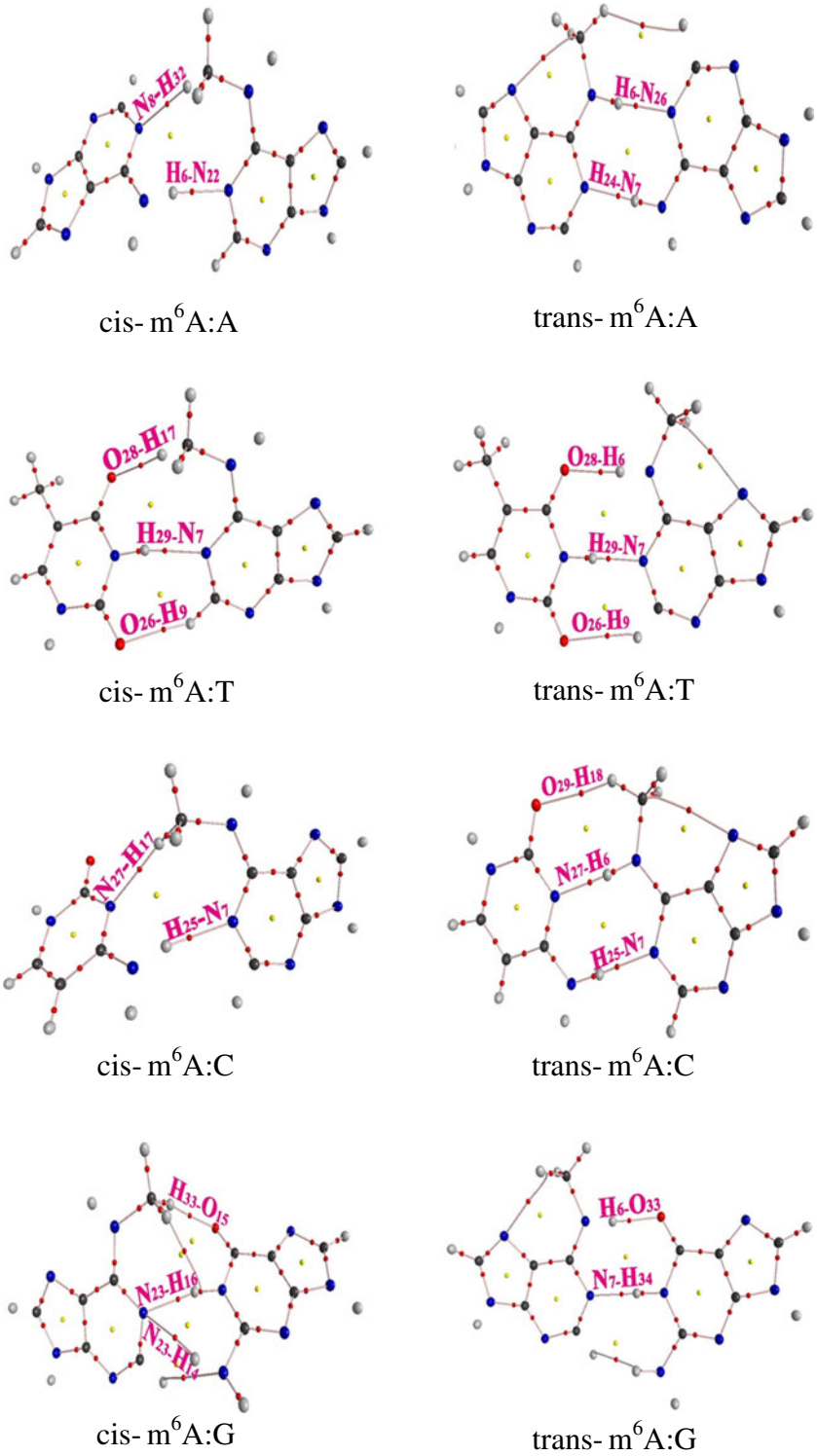


Fig. 5 AIM analysis (MP2/6-31G**, gas phase) of the base pairs with bond critical points, and the corresponding bond critical points are given

bonds contribute to the stability of base pair trans-m^{6}A:G and trans-m^{6}A:C. It should be pointed out that all the conclusions are fitted for the trends in aqueous solution.

Further investigation about the properties of the H-bonds, Wiberg bond indices (WBI) [51] were computed on the geometries with the MP2/6-31G** in gas phase and in aqueous solution and provided a way to judge the bond paths. The values of the WBI are also listed in Table 4. We can find that the WBI values agree well with the results of the AIM analysis, which reveal bond paths for the strong H-bond interactions. For example, the H-bond N_{24}-H_{29}…N_7 in geometry trans-m^{6}A:T (gas phase) has the biggest electron density (0.0412), and the biggest WBI (0.0775). The WBI values of the N-H…N/O

are much larger than C-H…N/O. The result further explains the N-H…N/O are much stronger than C-H…N/O.

From Tables 2, 3 and 4, Fig.4 and Fig.S3, the AIM results show that the base pairs cis-m^{6}A:T, cis-m^{6}A:G, trans-m^{6}A:T and trans-m^{6}A:C have three H-bonds while others only have two. We can see that in cis-m^{6}A:T, there is one N-H…N bond and two C-H…O bonds. It can be explained that a C-H…O in cis-m^{6}A:T pair weakens the stability of cis-m^{6}A:T conformer. Although trans-m^{6}A:T have two N-H…N/O bonds and one C-H…O bond, the H-bonds are not strong enough which can be deduced from the $E^{(2)}$, ρ_c, $\nabla^2\rho_c$ and WBI values. The pair trans-m^{6}A:G have only two H-bonds, fewer than that of the above base pairs, its stability is still the strongest of all the base pairs, it can be interpreted that interactions of H-bonds in trans-m^{6}A:G are the strongest of all, which make the complex more stable. It is interesting that in cis-m^{6}A:G pair, N_{12}-H_{14}…N_{23} is a weaker H-bond than C_{31}-H_{33}…O_{15}, because the H_{14}…N_{23} length is too long and the N_{12}-H_{14}…N_{23} angle is too small, thus weakening its strength. Therefore, the base pairs' stability is not directly related to the number of H-bonds, it significantly depends on the species and geometries which determine the H-bond properties.

Conclusions

In summary, MP2, AIM theory, NBO analysis and WBI have been employed to investigate the binding mechanism between m^{6}A and natural DNA bases in gas phase and in aqueous solution. The results show that the base pairs trans-m^{6}A:G and trans-m^{6}A:C have the most negative binding energies among the base pairs and are regarded as the most stable pairs. The H-bonds in aqueous solution are weaker than that in gas phase. This may explain that the solvent effect remarkably reduces the stability of the base pairs. It is also obvious that m^{6}A has a significant effect on the stability of base pairs. N6 site methylated decreases the binding energies compared to the normal Watson-Crick base pairs. In addition, the DNA bases are preferable to pair with trans-m^{6}A rather than cis-m^{6}A.

It is proven that these methods are the most efficient way for the characterization of the binding mechanism between m^{6}A and natural DNA bases. All of them implied that C-H…N/O contact should be classified as real but a rather weak H-bond. Besides, the type and geometry of H-bonds significantly influence the stability of base pairs. These calculations may be valuable to develop the property to study noncanonical base pairs, and improve the accuracy of DNA structure prediction.

Acknowledgments This work was supported by Doctor Candidate Foundation of Jiangnan University (JUDCF11006).

References

1. Watson JD, Crick FHC (1953) Molecular structure of nucleic acids. Nature 171(4356):737–738. doi:10.1038/171738a0
2. Szatyłowicz H, Sadlej-Sosnowska N (2010) Characterizing the strength of individual hydrogen bonds in DNA base pairs. J Chem Inf Model 50(12):2151–2161. doi:10.1021/ci100288h
3. Czyznikowska Z, Gora RW, Zalesny R, Lipkowski P, Jarzembska KN, Dominiak PM, Leszczynski J (2010) Structural variability and the nature of intermolecular interactions in watson−crick B-DNA base pairs. J Phys Chem B 114. doi:10.1021/jp101258q
4. Paragi G, Szajli E, Bogar F, Kovacs L, Guerra CF, Bickelhaupt FM (2008) Hydrogen bonding of 3- and 5-methyl-6-aminouracil with natural DNA bases. New J Chem 32(11):1981–1987. doi:10.1039/b803593h
5. Xue C, Popelier PLA (2008) Computational study of substituent effects on the interaction energies of hydrogen-bonded watson −crick cytosine:guanine base pairs. J Phys Chem B 112 (16):5257–5264. doi:10.1021/jp7108913
6. Xue C, Popelier PLA (2009) Prediction of interaction energies of substituted hydrogen-bonded watson−crick cytosine:Guanine8X base pairs. J Phys Chem B 113(10):3245–3250. doi:10.1021/jp8071926
7. Vanyushin BF, Tkacheva SG, Belozersky AN (1970) Rare bases in animal DNA. Nature 225(5236):948–949. doi:10.1038/225948a0
8. Ratel D, Ravanat JL, Berger F, Wion D (2006) N6-Methyladenine: the other methylated base of DNA. Bioessays 28(3):309–315. doi:10.1002/bies.20342
9. Wion D, Casadesús J (2006) N6-Methyl-adenine: an epigenetic signal for DNA–protein interactions. Nat Rev Microbiol 4(3):183–192. doi:10.1038/nrmicro1350
10. Barbeyron T, Kean K, Forterre P (1984) DNA adenine methylation of GATC sequences appeared recently in the Escherichia coli lineage. Journal of bacteriology 160 (2):586. doi:http://jb.asm.org/content/160/2/586.
11. Banerjee A, Rao DN (2011) Functional analysis of an acid adaptive DNA adenine methyltransferase from *helicobacter pylori* 26695. PLoS One 6(2):e16810. doi:10.1371/journal.pone.0016810
12. Robertson KD, Jones PA (2000) DNA methylation: past, present and future directions. Carcinogenesis 21(3):461–467. doi:10.1093/carcin/21.3.461
13. Chen RZ, Pettersson U, Beard C, Jackson-Grusby L, Jaenisch R (1998) DNA hypomethylation leads to elevated mutation rates. Nature 395(6697):89–93. doi:10.1038/25779
14. Vanyushin BF (2005) Enzymatic DNA methylation is an epigenetic control for genetic functions of the cell. Biochem Mosc 70 (5):488–499. doi:10.1007/s10541-005-0143-y
15. Vanyushin BF (2006) DNA methylation and epigenetics. Russ J Genet 42(9):985–997. doi:10.1134/s1022795406090055
16. Cheng X (1995) Structure and function of DNA methyltransferases. Annu Rev Biophys Biomol Struct 24(1):293–318. doi:10.1146/annurev.bb.24.060195.00 1453
17. Forde GK, Kedzierski P, Sokalski WA, Forde AE, Hill GA, Leszczynski J (2006) Physical nature of interactions within the active site of cytosine-5-methyltransferase. J Phys Chem A 110(6):2308–2313. doi:10.1021/jp056415u
18. Vanyushin BF, Belozersky AN, Kokurina NA, Kadirova DX (1968) 5-Methylcytosine and 6-methylaminopurine in bacterial DNA. Nature 218(5146):1066–1067. doi:10.1038/2181066a0
19. Collier J (2009) Epigenetic regulation of the bacterial cell cycle. Curr Opin Microbiol 12(6):722–729. doi:10.1016/j.mib.2009.08.005
20. Kamat SS, Fan H, Sauder JM, Burley SK, Shoichet BK, Sali A, Raushel FM (2011) Enzymatic deamination of the epigenetic base N-6-methyladenine. J Am Chem Soc 133:2080–2083. doi:org/10.1021/ja110157u

21. Low DA, Weyand NJ, Mahan MJ (2001) Roles of DNA adenine methylation in regulating bacterial gene expression and virulence. Infect Immun 69(12):7197–7204. doi:10.1128/IAI.69.12.7197-7204.2001
22. Vanyushin BF, Alexandrushkina NI, Kirnos MD (1988) N6-Methyladenine in mitochondrial DNA of higher plants. FEBS Lett 233(2):397–399. doi:10.1016/0014-5793(88)80469-1
23. Fedoreyeva LI, Vanyushin BF (2002) N6-Adenine DNA-methyltransferase in wheat seedlings. FEBS Lett 514(2–3):305–308. doi:10.1016/S0014-5793(02)02384-0
24. Li D, Delaney JC, Page CM, Yang X, Chen AS, Wong C, Drennan CL, Essigmann JM (2012) Exocyclic carbons adjacent to the N6 of adenine are targets for oxidation by the escherichia coli adaptive response protein AlkB. J Am Chem Soc. doi:10.1021/ja3010094
25. Engel J, von Hippel PH (1978) Effects of methylation on the stability of nucleic acid conformations. Studies at the polymer level. J Biol Chem 253(3):927-934. doi:http://www.jbc.org/content/253/3/927.short
26. Diekmann S (1987) DNA methylation can enhance or induce DNA curvature. The EMBO J 6(13):4213-4217. doi:http://www.ncbi.nlm.nih.gov/pmc/articles/PMC553906/
27. Lozynski M, Rusinska-Roszak D, Mack H-G (1998) Hydrogen bonding and density functional calculations: the B3LYP approach as the shortest Way to MP2 results. J Phys Chem A 102(17):2899–2903. doi:10.1021/jp973142x
28. Miertuš S, Scrocco E, Tomasi J (1981) Electrostatic interaction of a solute with a continuum. A direct utilizaion of AB initio molecular potentials for the prevision of solvent effects. Chem Phys 55 (1):117–129. doi:10.1016/0301-0104(81)85090-2
29. Miertus S, Tomasi J (1982) Approximate evaluations of the electrostatic free energy and internal energy changes in solution processes. Chem Phys 65(2):239–245
30. Hobza P, Zahradník R (1988) Intermolecular complexes: the role of van der Waals systems in physical chemistry and in the biodisciplines. Elsevier, Amsterdam
31. van Duijneveldt FB, van Duijneveldt-van de Rijdt JGCM, van Lenthe JH (1994) State of the art in counterpoise theory. Chem Rev 94(7):1873–1885. doi:10.1021/cr00031a007
32. Boys S, Bernardi F (1970) The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. Mol Phys 19(4):553–566. doi:10.1080/00268977000101561
33. Reed AE, Curtiss LA, Weinhold F (1988) Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. Chem Rev 88(6):899–926. doi:10.1021/cr00088a005
34. Bader RFW (1990) Atoms in molecules: a quantum theory. Oxford University Press, Oxford
35. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JAJ, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegaw J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2003) Gaussian 03. Gaussian Inc, Pittsburgh
36. Biegler-König F, Schönbohm J (2002) AIM2000, a program to analyze and visualize atoms in molecules, version 2.0, Bielefeld, Germany
37. Bae SH, Cheong HK, Cheong C, Kang S, Hwang DS, Choi BS (2003) Structure and dynamics of hemimethylated GATC sites. J Biol Chem 278(46):45987–45993. doi:10.1074/jbc.M306038200
38. Sternglanz H, Bugg CE (1973) Conformation of N6-methyladenine, a base involved in DNA modification: restriction processes. Science 182(4114):833. doi:10.1126/science.182.4114.833
39. Fazakerley GV, Quignard E, Teoule R, Guy A, Guschlbauer W (1987) A two-dimensional 1H-NMR study of the dam methylase site: comparison between the hemimethylated GATC sequence, its unmethylated analogue and a hemimethylated CATG sequence. Eur J Biochem 167(3):397–404. doi:10.1111/j.1432-1033.1987.tb13351.x
40. van de Sande J, Ramsing N, Germann M, Elhorst W, Kalisch B, von Kitzing E, Pon R, Clegg R, Jovin T (1988) Parallel stranded DNA. Science 241(4865):551–557. doi:10.1126/science. 3399890
41. Otto C, Thomas GA, Rippe K, Jovin TM, Peticolas WL (1991) The hydrogen-bonding structure in parallel-stranded duplex DNA is reverse Watson-Crick. Biochemistry 30(12):3062–3069. doi:10.1021/bi00226a012
42. Rentzeperis D, Rippe K, Jovin TM, Marky LA (1992) Calorimetric characterization of parallel-stranded DNA: stability, conformational flexibility, and ion binding. J Am Chem Soc 114(15):5926–5928. doi:10.1021/ja00041a003
43. Gorb L, Podolyan Y, Dziekonski P, Sokalski WA, Leszczynski J (2004) Double-proton transfer in adenine−thymine and guanine −cytosine base pairs. A post-hartree−fock ab initio study. J Am Chem Soc 126(32). doi:10119-10129. doi:10.1021/ja049155n
44. Reynisson J, Steenken S (2002) DFT studies on the pairing abilities of the one-electron reduced or oxidized adenine-thymine base pair. Phys Chem Chem Phys 4(21):5353–5358. doi:10.1039/B206342E
45. Guerra CF, Bickelhaupt FM, Snijders JG, Baerends EJ (1999) The nature of the hydrogen bond in DNA base pairs: the role of charge transfer and resonance assistance. Chemistry-Eur J 5:3581. doi:10.1002/(SICI)1521-3765(19991203)5:12<3581::AID-CHEM3581>3.0.CO;2-Y
46. Koch U, Popelier PLA (1995) Characterization of C-H-O hydrogen bonds on the basis of the charge density. J Phys Chem 99 (24):9747–9754. doi:10.1021/j100024a016
47. Bondi A (1964) van der Waals volumes and radii. J Phys Chem 68 (3):441–451. doi:10.1021/j100785a001
48. Bader RFW (1998) A bond path: a universal indicator of bonded interactions. J Phys Chem A 102(37):7314–7323. doi:10.1021/jp981794v
49. Popelier PLA (1998) Characterization of a dihydrogen bond on the basis of the electron density. J Phys Chem A 102(10):1873–1878. doi:10.1021/jp9805048
50. Parr RG, Yang W (1994) Density-functional theory of atoms and molecules, vol 16. Oxford University Press, USA
51. Wiberg K (1968) Application of the pople-santry-segal CNDO method to the cyclopropylcarbinyl and cyclobutyl cation and to bicyclobutane. Tetrahedron 24(3):1083–1096. doi:10.1016/0040-4020(68)88057-3