The effect of water on infrared spectra of DNA

2009 J. Phys.: Condens. Matter 21 064242

(http://iopscience.iop.org/0953-8984/21/6/064242)

J. Phys.: Condens. Matter **21** (2009) 064242 (4pp) doi:10.1088/0953-8984/21/6/064242

The effect of water on infrared spectra of DNA

Hisashi Taniguchi and Mineo Saito

Division of Mathematical and Physical Science, Graduate School of Natural Science and Technology, Kanazawa University, Kakuma-machi, Kanazawa 920-1192, Japan

E-mail: taniguchi@cphys.s.kanazawa-u.ac.jp

Received 18 July 2008, in final form 1 October 2008
Published 20 January 2009
Online at stacks.iop.org/JPhysCM/21/064242

Abstract

The effect of water on infrared (IR) spectra of DNA is studied by using first-principles calculations based on the hybrid density functional theory. Our calculations of frequencies of C=O stretching modes in DNA without water do not reproduce the IR spectra, whereas calculations for DNA surrounded by water reproduce the IR spectra well. We also find that the energy of adsorption of one water molecule around a C=O site is large ($\simeq$0.6 eV) owing to hydrogen bond formation. We therefore conclude that the effect of water plays an important role as regards IR spectra of DNA in a water solution.

1. Introduction

In the field of genomics, the focus is now shifting to the interplay between genes and proteins, and atomic level characterization of biomolecules is becoming more important. Infrared spectroscopy is expected to give useful information on biomolecules [1]. Recently, infrared absorption spectroscopy in the multiple-internal-reflection geometry (MIR-IRAS) in conjunction with electrophoresis has emerged as a tool of atomic level analysis of biomolecules in aqueous systems [2–7]. For example, by using MIR-IRAS, label-free detection of *in vitro* DNA was recently achieved [2, 3]. This technique enables detection of hybridization of DNA since the MIR-IRAS spectra of the double-strand (ds-)DNA and single-strand (ss-)DNA are clearly different. There are three main broad peaks at 1690, 1670 and 1640 cm^{-1} for ds-DNA, whereas the main peak of the sum of the two spectra of complementary ss-DNA appears at 1664 cm^{-1}.

Calculations of vibrational frequencies of DNA have already been carried out [2, 4, 8–12]. However, the study of the effect of water on the frequencies is still inadequate. In particular, the effect on the ss-DNA has not been revealed. Therefore, the difference between the effects of water on ds-DNA and ss-DNA has not been discussed on the basis of theoretical calculation.

In this paper, we perform first-principles calculations for DNA and clarify that the effects of water on the MIR-IRAS for the ss-DNA and ds-DNA are quite different. We find that the experimental MIR-IRAS are reproduced only when the effects of water are included. Our first-principles calculation suggests that the experimental main peaks mentioned above are due to C=O vibrations. When there is no effect of water, the vibrational peaks appear in the 1700–1800 cm^{-1} region in the case of ss-DNA whereas the peaks appear in the 1650–1750 cm^{-1} region for ds-DNA. These results are inconsistent with the experimental results. However, when the water molecules are attached to the systems, the spectra of ss-DNA and ds-DNA both show peaks in the region of 1600–1700 cm^{-1}, which is consistent with the experimental results. Therefore it is shown in this study that water molecules have an important role as regards the MIR-IRAS spectra.

2. Computational methods

In the *ab initio* molecular orbital theory based on cluster models, we employ the hybrid density functional theory as mentioned below. This method gives more accurate results than the local density approximation (LDA) and generalized gradient approximation (GGA). Its electron many-body energy is expressed as a mixture of the *exact* Hartree–Fock (HF) exchange and the DFT exchange–correlation functional [13]. We adopt Becke's three-parameter hybrid method using the Lee–Yang–Parr correlation functional (B3LYP):

$$F_{\mathrm{B3LYP}} = c^{\mathrm{X}}_{\mathrm{HF}} F^{\mathrm{X}}_{\mathrm{HF}} + (1 - c^{\mathrm{X}}_{\mathrm{HF}}) F^{\mathrm{X}}_{\mathrm{LDA}} + c^{\mathrm{X}}_{\mathrm{GGA}} \Delta F^{\mathrm{X}}_{\mathrm{B}} + c^{\mathrm{C}} F^{\mathrm{C}}_{\mathrm{VWN}} + (1 - c^{\mathrm{C}}) F^{\mathrm{C}}_{\mathrm{LYP}} \quad (1)$$

where F_{HF}, $F^{\mathrm{X}}_{\mathrm{LDA}}$, and $\Delta F^{\mathrm{X}}_{\mathrm{B}}$ denote the energies of the HF exchange, the LDA exchange, and its gradient correction due to Becke [14], respectively. As for the correlation, $F^{\mathrm{C}}_{\mathrm{VWN}}$,

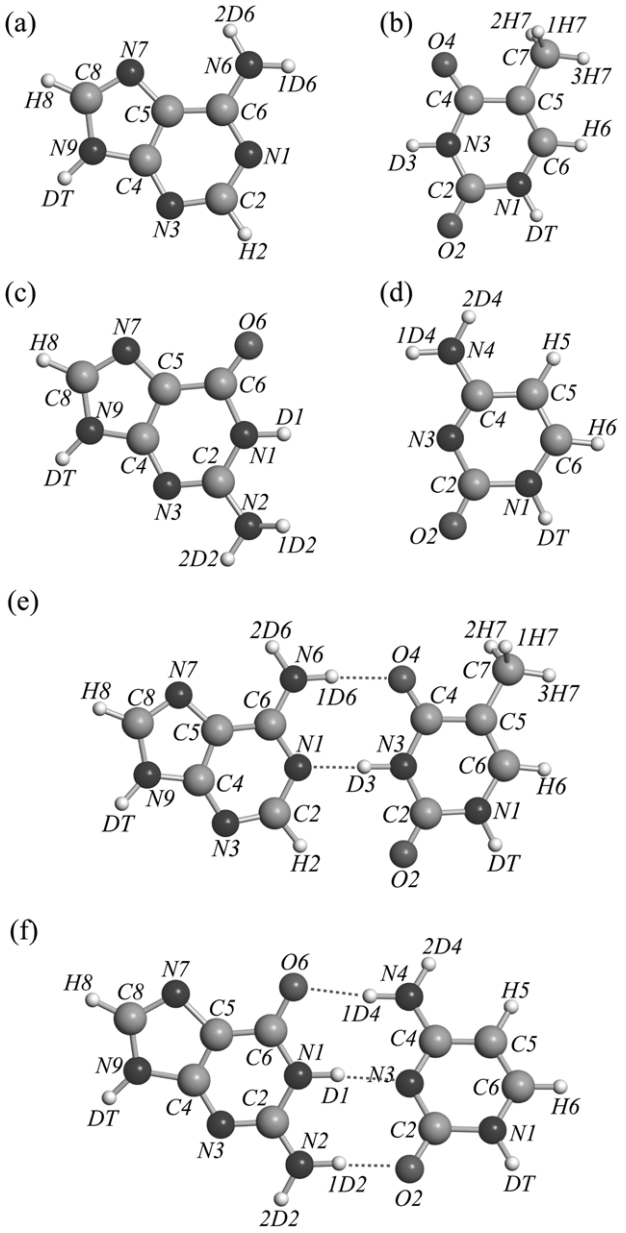

Figure 1. Atomic structures of (a) adenine, (b) thymine, (c) guanine, (d) cytosine, (e) an adenine–thymine pair, and (f) a guanine–cytosine pair. DT indicates the deuterium terminators which replace the sugar parts. Hydrogen atoms bonded to N and O atoms are replaced by deuteriums (see the text).

and F^{C}_{LYP} are the energies of the Vosko–Wilk–Nusair type [15] of correlation, and its gradient correction via LYP [16], respectively. Becke has determined the three parameters in (1) as $c^{x}_{HF} = 0.20$, $c^{X}_{GGA} = 0.72$ and $c^{C} = 0.19$ [13]. This hybrid scheme is considered to be a method going beyond the standard DFT method (LDA and GGA) and is highly reliable.

We use the linear combination of atomic orbital basis set labeled as 6-31G**. The core orbitals are expressed as linear combinations of the six-Gaussian basis set. The valence orbital splits into two independent orbitals in the variational calculation, and these orbitals consist of three and one primitive Gaussian functions, respectively. Further p (d) polarization of Gaussian orbitals augments the basis set for the hydrogen (the other) atoms.

Table 1. Difference of IR spectra for the DNA bases.

DNA base (C=O stretching mode)	Site	IR spectra without water molecule (cm^{-1})	IR spectra with water molecule (cm^{-1})
Thymine	O2	1759	1703
Thymine	O4	1713	1660
Guanine	O6	1754	1696
Cytosine	O2	1740	1674
Adenine–thymine pair	O2	1740	1688
Adenine–thymine pair	O4	1660	1653
Guanine–cytosine pair	O6–O2 (asym)	1690	1670
Guanine–cytosine pair	O6–O2 (sym)	1672	1634

The vibrational frequencies are estimated by calculations of the second derivatives of the total energy over the atomic coordinates. It is known that this harmonic oscillation tends to slightly overestimate vibrational frequencies, so we introduce so called scaling factor of 0.9614 [17]. We use cluster models where phosphoric acids bonded to DNA bases are replaced by deuterium (figure 1). Since MIR-IRAS experiments are performed for DNA in D_2O solution, hydrogen atoms bonded to N and O atoms in DNA are substituted by deuterium atoms. It is known that the hydrogen atoms bonded to carbon atoms are not substituted by deuterium [18]. We find that if the phosphoric acid is replaced by the deuterium atom, the calculated frequency varies within 10 cm^{-1} for most of the cases. In the case of thymine, the shift is about 30 cm^{-1}. All the calculations throughout this paper are performed by employing the Gaussian 03 program [19].

3. Results

We first study the three bases containing C=O bonds (thymine, guanine, and cytosine) neglecting the effect of water. The calculated frequencies for ss-DNA are shown in table 1. Thymine has two oxygen atoms labeled O2 and O4 (figure 1(b)). The frequencies of C=O bond stretching in the cases of the O2 and O4 atoms are 1759 cm^{-1} and 1713 cm^{-1}, respectively (table 1). The frequencies originating from the O6 atom in guanine (figure 1(c)) and the O2 atom in cytosine (figure 1(d)) are 1754 cm^{-1} and 1740 cm^{-1}, respectively (table 1). We conclude that the vibrational peaks due to C=O stretching modes appear in the region 1713–1759 cm^{-1}.

Next, we calculate the frequencies of the adenine–thymine (A–T) and guanine–cytosine (G–C) pairs. These two pairs mainly show peaks in the region 1660–1690 cm^{-1} and only one peak due to the A–T pair appears at 1740 cm^{-1} (table 1). In the G–C pair, two oxygen atoms (O2 in cytosine and O6 in guanine) are affected by hydrogen bonds (figure 1) and therefore the C=O bonds are weakened. As a result, these two bond stretching motions induce two vibrational peaks in the low frequency region (1672–1690 cm^{-1}). The higher (1690 cm^{-1}) and lower (1672 cm^{-1}) peaks are due to antisymmetric and symmetric motions of the two oxygen atoms. In the case of the A–T pair, the O4 in thymine

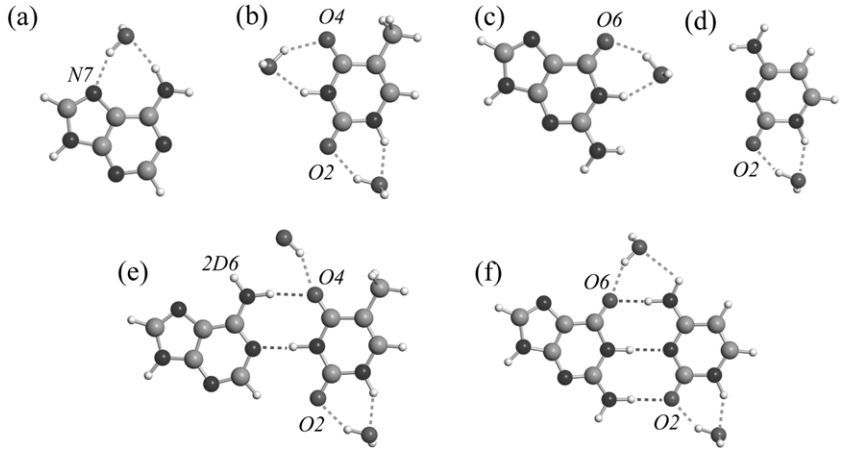

Figure 2. Water molecules are in a state of adsorption for the (a) adenine, (b) thymine, (c) guanine, (d) cytosine, (e) adenine–thymine pair and (f) guanine–cytosine pair.

is also affected by a hydrogen bond; thus the vibrational peak due to this oxygen atom appears in the low frequency region (1660 cm^{-1}) (figure 1(e) and table 1). The frequency (1740 cm^{-1}) originates from the C=O stretching mode of O2 in the A–T pair. This high frequency is due to the fact that the O2 atom does not form a hydrogen bond (figure 1(e)). We conclude that when there are no water molecules, the vibrational peaks of the ds-DNA appear in the region of 1660–1740 cm^{-1}.

The calculational peaks (1660–1690 cm^{-1}) of the ds-DNA are overall similar to the experimental peaks (1640, 1670 and 1690 cm^{-1}). However, there is one high frequency mode (1740 cm^{-1}) which is not seen in experimental spectra. Furthermore, the calculations for the ss-DNA are inconsistent with experimental results: experimental spectra of the ss-DNA have broad peaks at 1640 cm^{-1}, whereas the theoretical peaks appear in the region 1713–1759 cm^{-1}. These results strongly indicate that the effect of water on the vibrational spectra is not negligible.

The adsorption energies of the water molecules are tabulated in table 2. In the cases of thymine, guanine and cytosine, the adsorption energies per D_2O are found to be 0.6–0.7 eV. One D_2O forms two hydrogen bonds, i.e., two O···D bonds are formed (figure 2). Therefore, the formation energy of a single hydrogen bond is estimated to be 0.3–0.35 eV. The optimized hydrogen bond lengths are found to be 1.8–2.1 Å. For the pair of bases, the adsorption energies per D_2O are 0.5–0.6 eV, which are lower than those in the case of the bases in ss-DNA (0.6–0.7 eV). Because of the rather complicated geometries of the pairs, the hydrogen bonds originating from the adsorption of D_2O are weak. In particular, in the case of the adenine–thymine pair, one D_2O forms only one O···D bond and a hydrogen bond is not formed for the 2D6 atom in adenine (figure 2(e)).

In any case, the adsorption energies of D_2O are very large compared with the thermal energy at room temperature. Therefore, we expect the water molecules to be distributed near each DNA base.

We here briefly mention that the N atom also forms hydrogen bonds. In the case of adenine, a water molecule forms two hydrogen bonds, i.e., N···D and O···D bonds (figure 2(a)). The adsorption energy is 0.65 eV (table 2), which is considered to originate from the formation of O···D and N···D bonds. Since the formation energy of the O···D bond is estimated to be 0.3–0.35 eV, the formation energy of the N···D bond is expected to be similar to that of the O···D bond. Therefore N···D bonds as well as O···D bonds are expected to be formed when DNA is in D_2O. The formation of the N···D bond may affect the frequency of the ring deformation mode. However, we find that its effect on the vibrational frequency is within 10 cm^{-1}. Therefore, we neglect the formation of the N···D bonds in the calculation of vibrational frequencies.

Table 2. Energy of adsorption of the D_2O on the DNA bases.

	DNA base	D_2O (number)	Adsorption energy (eV)	Adsorption energy per D_2O (eV)
(a)	Adenine	1	0.645 066	0.645
(b)	Thymine	2	1.200 840	0.600
(c)	Guanine	1	0.693 283	0.693
(d)	Cytosine	1	0.698 739	0.699
(e)	Adenine–thymine pair	2	0.967 246	0.484
(f)	Guanine–cytosine pair	2	1.122 681	0.561

Then we perform first-principles calculations on DNA surrounded by water molecules to evaluate the effect of water on vibrational frequencies. When there is no effect of water, the vibrational peaks of ss-DNA appear in the region 1713–1759 cm^{-1} as was mentioned. On the other hand, when the effect of water is active, the vibrational peaks appear in the region 1660–1703 cm^{-1} (table 1). Therefore, we conclude that the effect of water substantially reduces the vibrational frequencies. As shown in figure 2, the O2 and O4 atoms in thymine, the O6 atom in guanine and the O2 atom in cytosine are affected by hydrogen bond formation. As a result, C=O bonds for these O atoms are weakened, reducing the vibrational frequencies of the C=O stretching mode.

For the cytosine–guanine pairs without water, peaks appear in the region of 1672–1690 cm^{-1} (table 1). These rather low frequencies are due to the fact that hydrogen bonds are

formed. When water is attached to the O atoms, the frequencies are only slightly lowered (1670–1634 cm^{-1}; table 1). Since O2 and O6 atoms already form hydrogen bonds when there is no water, the effect of water on the frequency is very small. For the adenine–thymine pair without water molecules, one peak due to the O2 atom appears at 1740 cm^{-1}. This high frequency is due to the fact that the O2 atom in thymine does not have a hydrogen bond. When the water molecule is bonded to this atom, the frequency is substantially lowered (1688 cm^{-1}). The vibrational frequency due to the O4 atom in thymine is low (1660 cm^{-1}) when there is no water. When the water is bonded to the O4 atom, the frequency is slightly changed (1653 cm^{-1}). Therefore only one frequency due to the O2 atom in thymine among the four vibrational frequencies in the pairs is substantially lowered by the effect of water.

As mentioned above, when the water is bonded to the O atoms, the absorption peak due to the C=O stretching modes appears in the region 1600–1700 cm^{-1} in both cases, ss-DNA and ds-DNA. These results are consistent with the experimental results. Therefore, we conclude that the water plays a significant role as regards the absorption spectrum.

4. Summary

We perform first-principles calculations for vibrational frequencies in DNA. Experimental spectra are well reproduced when the effects of water are included. On the other hand, if we do not consider the effect of water, the experimental spectra are not reproduced. Therefore, we conclude that the water plays an important role as regards infrared absorption spectra. Our calculation shows that the frequencies of the C=O stretching modes are reduced when O has a hydrogen bond. Therefore, when there is no water, the frequencies of the ds-DNA are in general lower than those of the ss-DNA since ds-DNA has some hydrogen bonds due to hybridization. When there are water molecules around the DNA, hydrogen bonds are newly formed. Compared with the frequencies of the ds-DNA, the frequencies of the ss-DNA are substantially reduced by the effect of water; then the experimental spectra are well reproduced. The substantial reduction of frequencies in the case of the ss-DNA is due to the fact that the ss-DNA has no hydrogen bond when there is no water.

Acknowledgments

This work was supported by a Grant-in-Aid for Scientific Research (No. 17510097) from the Japan Society for the Promotion of Science. The computation in this work was performed using the facilities of the Supercomputer Center, Institute for Solid State Physics, University of Tokyo. The authors thank the Research Center for Computational Science, Okazaki Research Facilities, National Institutes of Natural Sciences, for the use of the facilities.

References

[1] For an example, see Pevsner A and Diem M 2001 *Appl. Spectrosc.* **55** 1502–5
[2] Miyamoto K, Ishibashi K, Yamaguchi R, Kimura Y, Ishii H and Niwano M 2006 *J. Appl. Phys.* **99** 094702
[3] Miyamoto K, Ishibashi K, Hiroi K, Kimura Y, Ishii H and Niwano M 2005 *Appl. Phys. Lett.* **86** 053902
[4] Miyamoto K, Ishibashi K, Yamaguchi R, Kimura Y, Ishii H and Niwano M 2005 *Surf. Sci.* **26** 553
[5] Shiraki H, Kimura Y, Ishii H, Ono S, Itaya K and Niwano M 2004 *Appl. Surf. Sci.* **237** 369–73
[6] Kimura Y, Nemoto J, Shinohara M and Niwano M 2003 *Phys. Status Solidi* a **197** 577–81
[7] Kimura Y, Kondo Y and Niwano M 2001 *Appl. Surf. Sci.* **175/176** 157
[8] Krummel A T and Zanni M T 2006 *J. Phys. Chem.* B **110** 13991–4000
[9] Lee C, Park K H and Cho M 2006 *J. Chem. Phys.* **125** 114508
[10] Lee C and Cho M 2006 *J. Chem. Phys.* **125** 114509
[11] Lee C, Park K H, Kim J A, Hahn S and Cho M 2006 *J. Chem. Phys.* **125** 114510
[12] Lee C and Cho M 2007 *J. Chem. Phys.* **126** 145102
[13] Becke A D 1993 *J. Chem. Phys.* **98** 5648
[14] Becke A D 1988 *Phys. Rev.* A **38** 3098
[15] Vosko S H, Wilk L and Nusair M 1980 *Can. J. Phys.* **58** 1200
[16] Lee C, Yang W and Parr R G 1988 *Phys. Rev.* B **37** 785
[17] Scott A P and Radom L 1996 *J. Phys. Chem.* **100** 16502
[18] Duchesne J (ed) 1973 *Physico-Chemical Properties of Nucleic Acids* (London: Academic)
[19] Frisch M J, Trucks G W, Schlegel H B, Scuseria G E, Robb M A, Cheeseman J R, Montgomery J A Jr, Vreven T, Kudin K N, Burant J C, Millam J M, Iyengar S S, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson G A, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox J E, Hratchian H P, Cross J B, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann R E, Yazyev O, Austin A J, Cammi R, Pomelli C, Ochterski J W, Ayala P Y, Morokuma K, Voth G A, Salvador P, Dannenberg J J, Zakrzewski V G, Dapprich S, Daniels A D, Strain M C, Farkas O, Malick D K, Rabuck A D, Raghavachari K, Foresman J B, Ortiz J V, Cui Q, Baboul A G, Clifford S, Cioslowski J, Stefanov B B, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin R L, Fox D J, Keith T, Al-Laham M A, Peng C Y, Nanayakkara A, Challacombe M, Gill P M W, Johnson B, Chen W, Wong M W, Gonzalez C and Pople J A 2004 *Gaussian 03, Revision C.02* (Wallingford, CT: Gaussian)