Theor Chem Acc (2011) 130:385–391
DOI 10.1007/s00214-011-1004-y

REGULAR ARTICLE

Nuclear quantum effect on the hydrogen-bonded structure of guanine–cytosine pair

Masashi Daido · Akihito Koizumi · Motoyuki Shiga · Masanori Tachikawa

Received: 29 April 2011 / Accepted: 14 July 2011 / Published online: 4 August 2011

Abstract The structure of Watson–Crick type guanine–cytosine (G–C) base pair has been studied by classical hybrid Monte Carlo (HMC) and quantum path integral hybrid Monte Carlo (PIHMC) simulations on the semi-empirical PM6 potential energy surface. For the three NH···X hydrogen-bonded moieties, the intramolecular NH bonds are found systematically longer while the H···X distance shorter in the PIHMC simulation than in the HMC simulation. We found that the hydrogen bonded length N···X correlates with the H···X distance, but not with the NH distance. A correlation is also between the neighboring hydrogen bonds in the G–C base pair.

Keywords Watson–Crick type guanine–cytosine base pair · Path integral simulation · Nuclear quantum effect

1 Introduction

Hydrogen bonding between nucleobase pairs is a basic molecular interaction that constitutes the double stranded structure of DNA [1]. Over the decades, many efforts have been devoted to obtain the detailed structure of Watson–Crick type nucleobase pairs [2] as well as their isomers using experimental techniques and theoretical approaches. Experimentally, vibrational spectroscopy has provided very useful information about isolated nucleobase molecules and the hydrogen-bonded structure of nucleobase pairs [3–12]. For instance, Choi et al. [11] have detected the signal of four different tautomers of isolated guanine by infrared laser spectroscopy in helium nanodroplets. Abo-Riziq et al. [12] have been able to distinguish various isomers of guanine–cytosine (G–C) base pairs by IR–UV hole burning spectra including the Watson–Crick type structure. In these studies, the vibrational bands observed have been assigned by comparing them with those obtained from ab initio calculations.

So far, most of the ab initio calculations of nucleobase pairs have been carried out by simple normal mode analysis [13–22] and classical molecular dynamics (MD) methods [23–27]. However, the role of quantum effects of vibration in addition to anharmonic and mode coupling effects could be important for more accurate description of hydrogen-bonded systems. It has been evident from the calculations based on vibrational theories [28, 29] that these effects are indispensible to understand the complex behavior of the infrared spectrum of G–C base pair. Recently, Pérez et al. [30] have suggested that nuclear quantum effect can change the relative stability of tautomeric forms of Watson–Crick base pair models from Car–Parrinello path integral molecular dynamics simulation [31, 32]. To our knowledge, there is no publication about the structure of full G–C base pair at finite temperature taking account of the nuclear quantum effect.

In this study, we report hybrid Monte Carlo and path integral hybrid Monte Carlo simulations [33] of Watson–

Dedicated to Professor Shigeru Nagase on the occasion of his 65th birthday and published as part of the Nagase Festschrift Issue.

Electronic supplementary material The online version of this article (doi:10.1007/s00214-011-1004-y) contains supplementary material, which is available to authorized users.

M. Daido · A. Koizumi · M. Tachikawa (✉)
Quantum Chemistry Division, Graduate School of Science, Yokohama-city University, Seto 22-2, Kanazawa-ku, Yokohama 236-0027, Japan
e-mail: tachi@yokohama-cu.ac.jp

M. Shiga
CCSE, Japan Atomic Energy Agency (JAEA), Higashi-Ueno 6-9-3, Taito-ku, Tokyo 110-0015, Japan

Crick type G–C base pair in the gas phase (Fig. 1). It has been reported that Watson–Crick type G–C base pair is the most stable structure within the MP2/TZVPP calculation [12]. It should be noted that this form is different from the keto-guanine enol-cytosine type G–C pair [28, 29] whose vibrational spectrum is studied well by experiments as well as vibrational theories. While Pérez et al. [30] use density function theory for the electronic structure, we employ semiempirical PM6 method that is less accurate. However, the present simulations are carried out for the whole G–C system including the purine and pyrimidine rings to guarantee proper description of the planar structure of G–C base pair. Here, we mainly discuss the difference between the classical and quantum descriptions of the three N···X hydrogen bonds in the G–C base pair.

2 Computational method

All the electronic structure calculations are based on semiempirical PM6 method. In order to confirm the validity of PM6, we have performed preliminary calculations with various ab initio computational levels for the isolated G–C base pair. Also, we have shown the vibrational frequency data obtained from normal mode analyses in Supporting information, together with the measured and calculated reference values. From Table S1 in Supporting information, it is found that PM6 tends to underestimate the vibrational frequencies of G–C base pair as compared to the other ab initio methods. The parameters characterizing three hydrogen bonds of the geometry-optimized G–C base pair are listed in Table 1. (The atoms are labeled as in Fig. 1.) The results are shown for HF/6-31G**, MP2/6-31G**, B3LYP/6-31G**, and PM6 calculations. As for the covalent-bonded NH distances, the values of PM6 are close to those of B3LYP and MP2, while they are slightly longer than the HF results. The values of the hydrogen bonded distances of PM6 are in between the values of HF and MP2. The hydrogen bonded distances of PM6 are $R_{N1\cdots O1} > R_{N2\cdots N'2} > R_{N3\cdots O3}$ similar to the case of MP2. From Table 1, thus, we find that PM6 is qualitatively sufficient at least to describe the stable geometry of the hydrogen bonds of G–C base pair.

The hybrid Monte Carlo simulation (HMC) and path integral hybrid Monte Carlo (PIHMC) simulation have been carried out in the same way as in the previous work [33] at 300 K. The PIHMC simulation (or "quantum simulation") has been performed for 400,000 steps with the number of beads $P = 16$, step size $\Delta t = 1.5$ fs. The HMC simulation (or "classical simulation") has been executed for 1,920,000 steps with $\Delta t = 1.8$ fs. For the sake of sampling efficiency, we have chosen Δt so as the acceptance ratios are around 75–80% from short HMC/PIHMC runs. The final acceptance ratios were 77 and 75% for the quantum and classical simulations, respectively.

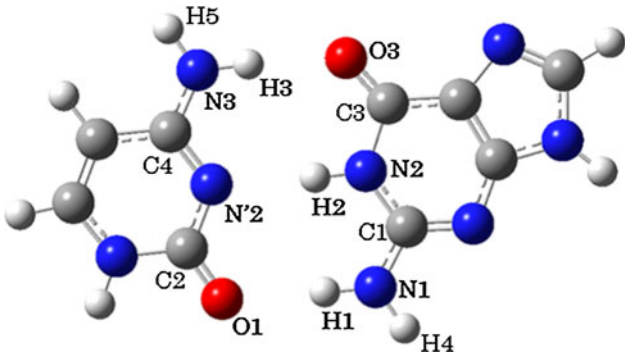

Fig. 1 Schematic illustration of guanine–cytosine base pair

3 Results and discussion

3.1 One-dimensional analysis

First, let us look at the three NH covalent bonds. The distributions of the covalent bond lengths R_{N1H1}, R_{N2H2}, and R_{N3H3} obtained from the quantum and classical simulations, together with the equilibrium NH bond lengths, are shown in Fig. 2a–c, where the superscripts (qm), (cl), and

Table 1 Structural parameters of three hydrogen bonds in G–C base pair with HF/6-31G**, MP2/6-31G**, B3LYP/6-31G**, and PM6 level of calculations

	HF/6-31G**	MP2/6-31G**	B3LYP/6-31G**	PM6
R_{N1H1}	1.001	1.018	1.022	1.028
R_{N2H2}	1.007	1.032	1.034	1.049
R_{N3H3}	1.007	1.029	1.038	1.039
$R_{H1\cdots O1}$	2.018	1.929	1.904	1.981
$R_{H2\cdots N'2}$	2.033	1.906	1.896	1.955
$R_{H3\cdots O3}$	1.919	1.787	1.749	1.868
$R_{N1\cdots O1}$	3.019	2.945	2.926	3.008
$R_{N2\cdots N'2}$	3.039	2.935	2.929	2.985
$R_{N3\cdots O3}$	2.925	2.816	2.787	2.905

Unit in Å

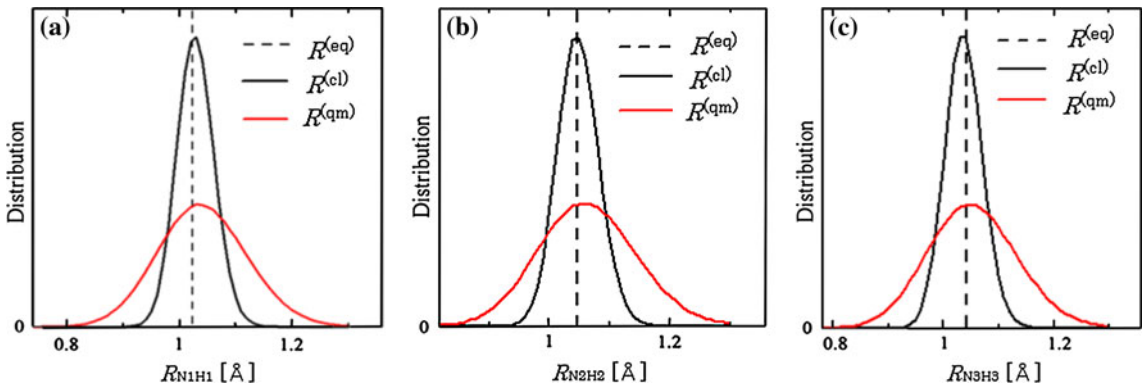

Fig. 2 One-dimensional distributions of R_{NH} on **a** R_{N1H1}, **b** R_{N2H2}, and **c** R_{N3H3} with quantum and classical simulations, as well as the equilibrium values, respectively

Table 2 Average values and statistical errors of the covalent bond lengths (R_{NH}) with quantum and classical simulations, as well as the equilibrium values

	Quantum	Classical	Equilibrium
R_{N1H1}	1.041 ± 0.000	1.028 ± 0.000	1.028
R_{N2H2}	1.065 ± 0.001	1.048 ± 0.000	1.049
R_{N3H3}	1.054 ± 0.001	1.037 ± 0.000	1.039

Unit in Å

Table 3 Average values and statistical errors of hydrogen bonded distances ($R_{N\cdots X}$) with quantum and classical simulations, as well as the equilibrium values

	Quantum	Classical	Equilibrium
$R_{N1\cdots O1}$	3.051 ± 0.014	3.109 ± 0.016	3.008
$R_{N2\cdots N'2}$	2.993 ± 0.012	3.050 ± 0.006	2.985
$R_{N3\cdots O3}$	2.958 ± 0.023	3.004 ± 0.013	2.905

Unit in Å

(eq) stand for quantum, classical, and the equilibrium structure, respectively. The average values and the statistical errors of the covalent bond lengths are listed in Table 2. We can see that the R_{NH} distributions in the classical simulation have sharp peaks around the equilibrium (EQ) values, while those in quantum simulation are more broadened due to the zero-point motion. Thus, the average NH bond lengths of the classical simulation are very close to the values of the EQ structure, while the average NH bond lengths of the quantum simulation are longer. We can confirm the values in Table 2 that $R_{NH}^{(eq)} \approx \langle R_{NH}^{(cl)} \rangle < \langle R_{NH}^{(qm)} \rangle$ with respect to all the three NH bonds. This result shows that nuclear quantum effect has the dominant contribution to the anharmonicity of the covalent NH bond at 300 K.

Next, the distributions of the hydrogen bonded distances measured between two heavy atoms ($R_{N\cdots X}$ where X = O or N′) are shown in Fig. 3a–c. The average values and the statistical errors are listed in Table 3. Figure 3a–c indicates that the peak positions of $R_{N\cdots O}$ and $R_{N\cdots N'}$ in the quantum simulation are found to be close to the corresponding values of the equilibrium structure, while they are longer in the classical simulation. We can confirm the average values in Table 3 that $R_{N\cdots X}^{(eq)} \approx \langle R_{N\cdots X}^{(qm)} \rangle < \langle R_{N\cdots X}^{(cl)} \rangle$ for all the three

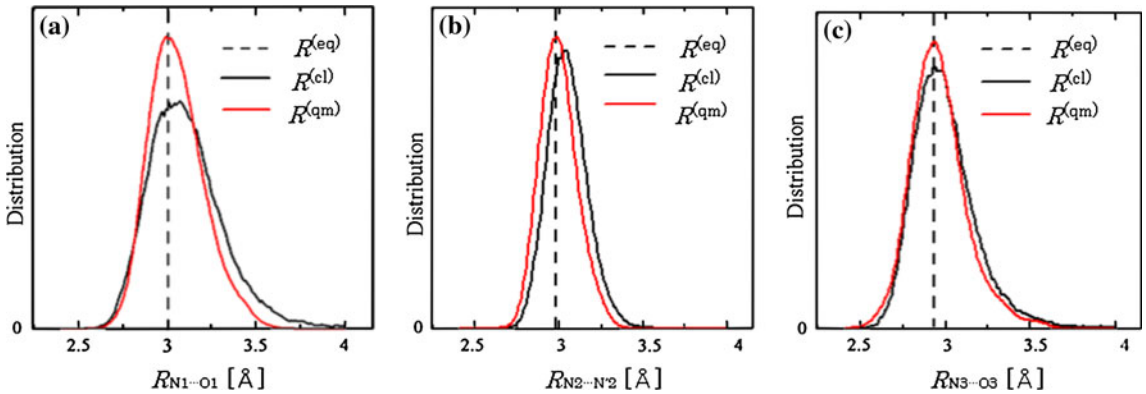

Fig. 3 One-dimensional distributions of $R_{N\cdots X}$ on **a** $R_{N1\cdots O1}$, **b** $R_{N2\cdots N'2}$, and **c** $R_{N3\cdots O3}$ with quantum and classical simulations, as well as the equilibrium values, respectively

Table 4 Average values and statistical errors of H···X distances with quantum and classical simulations, as well as the equilibrium values

	Quantum	Classical	Equilibrium
$R_{H1\cdots O1}$	2.067 ± 0.017	2.133 ± 0.018	1.981
$R_{H2\cdots N'2}$	1.987 ± 0.017	2.067 ± 0.009	1.955
$R_{H3\cdots O3}$	1.969 ± 0.023	2.030 ± 0.016	1.868

Unit in Å

hydrogen bonds. We also find that the quantum distributions of $R_{N\cdots N'}$ and $R_{N\cdots O}$ are slightly more localized than classical distribution. This is in contrast to the case of the covalent NH bond, for which quantum distributions are more delocalized than the classical counterpart as shown in Fig. 2a–c.

The average values and the statistical errors of $R_{H\cdots X}$ are shown in Table 4. By comparing between the average H···X distances of quantum and classical simulations, we find H···X distance shortened by nuclear quantum effect,

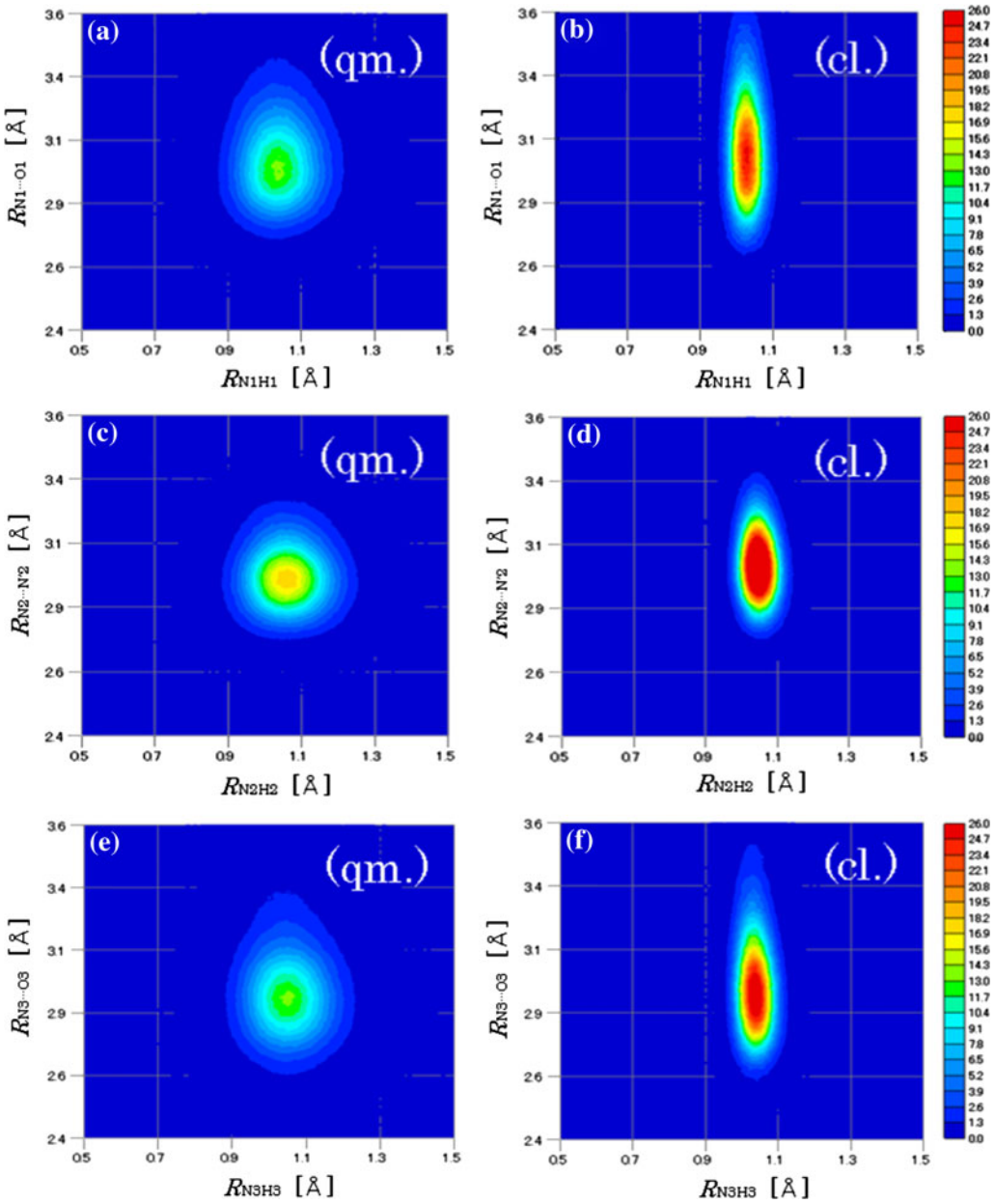

Fig. 4 Two-dimensional distributions of R_{NH} and $R_{N\cdots X}$ on **a** quantum R_{N1H1} and $R_{N1\cdots O1}$, **b** classical R_{N1H1} and $R_{N1\cdots O1}$, **c** quantum R_{N2H2} and $R_{N2\cdots N'2}$, **d** classical R_{N2H2} and $R_{N2\cdots N'2}$, **e** quantum R_{N3H3} and $R_{N3\cdots O3}$, and **f** classical R_{N3H3} and $R_{N3\cdots O3}$, respectively

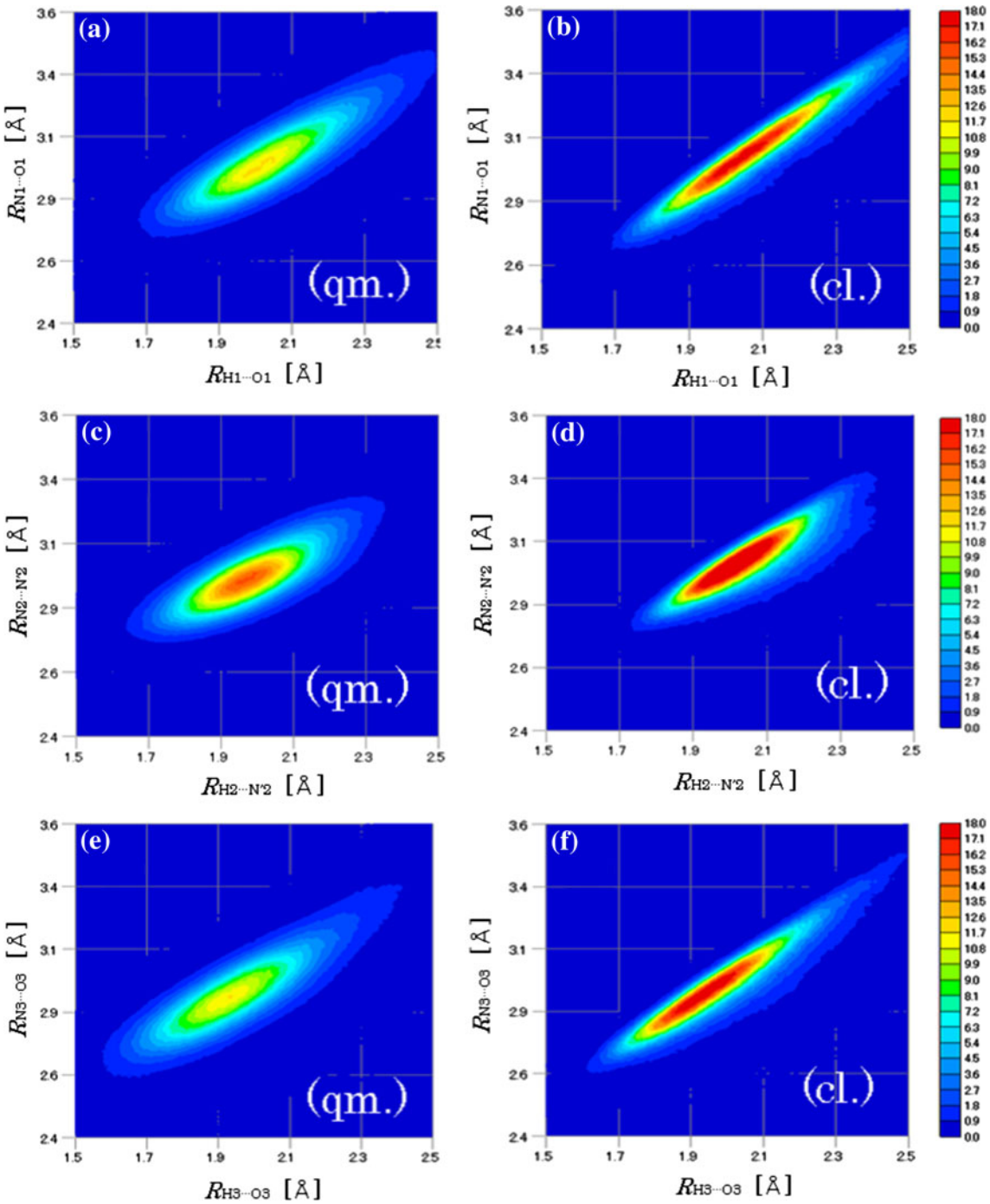

Fig. 5 Two-dimensional distributions of $R_{H\cdots X}$ and $R_{N\cdots X}$ on (**a**) quantum $R_{H1\cdots O1}$ and $R_{N1\cdots O1}$, (**b**) classical $R_{H1\cdots O1}$ and $R_{N1\cdots O1}$, (**c**) quantum $R_{H2\cdots N'2}$ and $R_{N2\cdots N'2}$, (**d**) classical $R_{H2\cdots N'2}$ and $R_{N2\cdots N'2}$, (**e**) quantum $R_{H3\cdots O3}$ and $R_{N3\cdots O3}$, and (**f**) classical $R_{H3\cdots O3}$ and $R_{N3\cdots O3}$, respectively

which is in contrast to the case of the covalent NH bond. The amount of shortening of the average H···X distance (0.07, 0.08, 0.06 Å for the respective hydrogen bonds) overwhelms the amount of lengthening of the average NH distance (0.013, 0.017, 0.017 Å). This is why the average N···X distance is shortened by nuclear quantum effect as shown in Table 3.

3.2 Two-dimensional analysis

Now, it would be interesting to see whether there is a correlation between the hydrogen bonded length and the position of intermediate hydrogen, as is done in the previous studies for various hydrogen-bonded systems [33–36]. We plot in Figs. 4 and 5 the two-dimensional

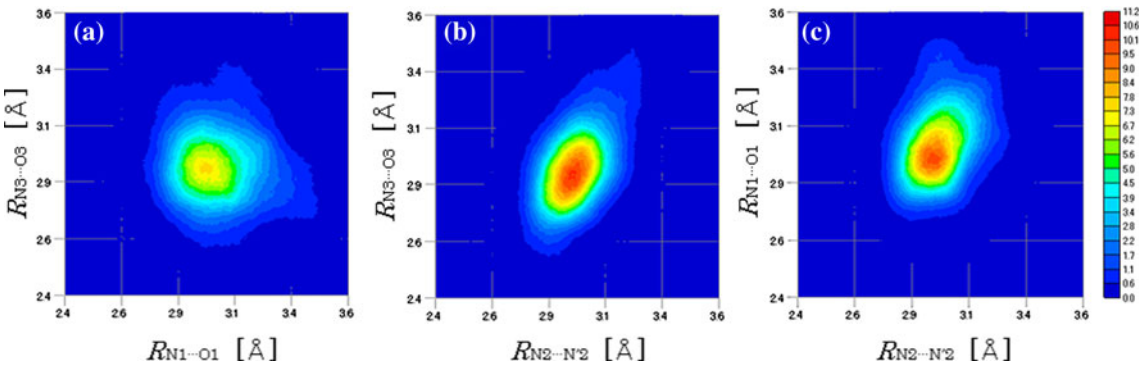

Fig. 6 Two-dimensional distributions of $R_{N\cdots X}$ and $R_{N\cdots X}$ on (**a**) $R_{N1\cdots O1}$ and $R_{N3\cdots O3}$, (**b**) $R_{N2\cdots N'2}$ and $R_{N3\cdots O3}$, and (**c**) $R_{N2\cdots N'2}$ and $R_{N1\cdots O1}$, respectively. Note that these are results with the quantum simulation, and almost the same tendency can be seen with the classical simulation

distributions with respect to the distances R_{NH} and $R_{N\cdots X}$ and the distances $R_{H\cdots X}$ and $R_{N\cdots X}$, respectively, for the three hydrogen bonds from the quantum and classical simulations. The shapes of two-dimensional distributions in Fig. 4 indicate the correlation between the NH and $N\cdots X$ distances is quite weak for both the classical and quantum simulations. For instance, as we have looked closer into Fig. 4a, one-dimensional NH distributions cut at the line $N\cdots X = 2.8$ Å and at the line $N\cdots X = 3.3$ Å have almost perfectly matched with each other (See in Supporting information). Therefore, we conclude that the $N\cdots X$ distance strongly correlates with the $H\cdots X$ distance as we see in Fig. 5, but not so much with the NH distance in Fig. 4.

We note in passing that the distributions in Fig. 5c, d are narrower than those of Fig. 5a, b, e, f, which implies that the $N2\cdots N'2$ hydrogen bond is tightly bonded than the other two $N\cdots X$ bonds.

3.3 The correlations among hydrogen bonds

Finally, Fig. 6 shows the two-dimensional distribution with respect to the two heavy-atom distances out of the three hydrogen bonds. Figure 6b, c indicates that there are correlations between the neighboring hydrogen bonds, i.e., between $R_{N2\cdots N'2}$ and $R_{N1\cdots O1}$, and between $R_{N2\cdots N'2}$ and $R_{N3\cdots O3}$. When one of the hydrogen bonds is short (long), the other tends to be also short (long). From Fig. 6a, it seems that the correlation between distant hydrogen bonds, i.e. between $R_{N1\cdots O1}$ and $R_{N3\cdots O3}$, is much weaker.

4 Conclusions

We have carried out classical hybrid Monte Carlo and quantum path integral hybrid Monte Carlo simulations of Watson–Crick type guanine–cytosine base pair in the gas phase using semiempirical PM6 potential energy surface. In the classical simulation that only includes thermal contribution to anharmonic vibration, the average covalent NH bond lengths have not changed much from the equilibrium bond lengths. The average NH bond lengths are found longer in the quantum simulation where both the quantum and thermal contributions to anharmonic vibration are included. The average hydrogen bonded distances measured between two heavy atoms $N\cdots X$ have been found shorter in the quantum simulation than those in the classical counterpart. This is presumably because the $N\cdots X$ distance is correlated with $H\cdots X$ distance, but not with NH distance. The correlation is also found with respect to the neighboring hydrogen bonded lengths in the G–C base pair. As a future work, ab initio path integral simulations of G–C base pair would be of interest to confirm the present results.

Acknowledgments The authors thank Prof. K. Yagi at University of Illinois for fruitful discussion. We would like to acknowledge Grant-in-Aid for Scientific Research and for the priority area by Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

1. Sutor J (1962) Nature 195:68–69
2. Watson J, Crick HC (1953) Nature 171:737–738
3. Bakker MJ, Compagnon I, Meijer G, Helden G, Kabelac M, Hobza P, Vries SM (2004) Phys Chem Chem Phys 6:2810–2815
4. Sobolewski AL, Domcke W, Hättig C (2005) Proc Natl Acad Sci 102:17903–17906
5. Abo-Riziq A, Crews B, Grace L, Vries SM (2005) J Am Chem Soc 127:2374–2375
6. Nir E, Janzen C, Imhof P, Kleinermanns K, Vries SM (2002) Phys Chem Chem Phys 4:732–739
7. Belau L, Wilson RK, Leone RS, Ahmed M (2007) J Phys Chem A 111:7562–7568
8. Pervushin K, Ono A, Fernández C, Szyperski T, Kainosho M, Wüthrich K (1998) Proc Natl Acad Sci 95:14147–14151

9. Hobza P, Spirko V (2003) Phys Chem Chem Phys 5:1290
10. Urashima S, Asami H, Ohba M, Saigusa H (2010) J Phys Chem A 114:11231–11237
11. Choi MY, Miller RE (2006) J Am Chem Soc 128:7320
12. Abo-Riziq A, Grace L, Nir E, Kabelac M, Hobza P, Vries SM (2005) Proc Natl Acad Sci 102:20
13. Norinder U (1987) J Mol Struct 151:259–269
14. Jorgensen LW, Pranata J (1990) J Am Chem Soc 112:2008–2010
15. Sponer J, Jurecka P, Hobza P (2004) J Am Chem Soc 126:10142–10151
16. Russo N, Toscano M, Grand A, Jolibois F (1998) J Comput Chem 19:989–1000
17. Douhal A, Guallar V, Moreno M, Lluch MJ (1996) Chem Phys Let 256:370–376
18. Gould RI, Kollman AP (1994) J Am Chem Soc 116:2493–2499
19. Sobolewski AL, Domcke W (2004) Phys Chem Chem Phys 6:2763–2771
20. Noguera M, Sodupe M, Bertran J (2004) Theor Chem Acc 112:318–326
21. Gorb L, Podolyan Y, Dziekonski P, Sokalski WA, Leszczynski J (2004) J Am Chem Soc 126:10119–10129
22. Shishkin OV, Sponer J, Hobza P (1999) J Mol Struct 477:15–21
23. Kabelac M, Hobza P (2001) J Phys Chem B 105:5804–5817
24. Zoete V, Meuwly M (2004) J Chem Phys 121:4377–4388
25. Villani G (2006) Chem Phys 324:438–446
26. Tsui V, Case AD (2000) J Am Chem Soc 122:2489–2498
27. Maranón J, Fantoni A, Grigera RJ (1999) J Theor Biol 201:93–102
28. Brauer B, Gerber RB, Kabelac M, Hobza P, Bakker MJ, Abo-Riziq A, Vries SM (2005) J Phys Chem A 109:6974–6984
29. Yagi K, Karasawa H, Hirata S, Hirao K (2009) Chem Phys Chem 10:1442–1444
30. Pérez A, Tuckerman ME, Hjalmarson PH, Lilienfeld AO (2010) J Am Chem Soc 132:11510–11515
31. Marx D, Parrinello M (1996) J Chem Phys 104:4077–4082
32. Tuckerman ME, Marx D, Klein ML, Parrinello M (1996) J Chem Phys 104:5579–5588
33. Suzuki K, Tachikawa M, Shiga M (2010) J Chem Phys 132:144108
34. Koizumi A, Suzuki K, Shiga M, Tachikawa M (2011) J Chem Phys 134:031101
35. Tachikawa M, Shiga M (2005) J Am Chem Soc 127:11908–11909
36. Suzuki K, Shiga M, Tachikawa M (2008) J Chem Phys 129:144310