

## The Effect of Water in Four Adenine-Thymine and Three Guanine-Cytosine Pairs: Combining Quantum and Statistical Mechanics

Jinkeong Lee and Sihyun Ham

*Department of Chemistry, Sookmyung Woman's University  
Cheongpa-ro-47-gil 100, Yongsan-Ku, Seoul, 140-742, Korea*

**E-mail:** sihyun@sookmyung.ac.kr

The molecular interactions between the nucleic acid bases and water molecules are important in organism. Despite Adenine-Thymine Hoogsteen base pair and Guanine-Cytosine Watson-Crick base pair have been demonstrated to be most stable in a gas phase, the effect of water on the stability of these base pairs remains elusive. Here we report the structural and thermodynamic characteristics on possible Adenine-Thymine and Guanine-Cytosine base pairs in a gas phase as well as in an aqueous phase by using quantum mechanical method and statistical mechanical calculations. First, we optimized the direct base-pair interaction energies of four Adenine-Thymine base pairs (Hoogsteen base pair, reverse Hoogsteen base pair, Watson-Crick base pair, and reverse Watson-Crick base pair) and three Guanine-Cytosine base pairs (GC1 base pair, GC2 base pair, and Watson-Crick base pair) in a gas phase at the B3LYP/6-31+G\*\* level. Then, the effect of solvent was quantified by the electronic reorganization energy and the solvation free energy by statistical mechanical calculations. Thereby, we discuss the effect of water on the stability of Adenine-Thymine and Guanine-Cytosine base pairs, and argue why Adenine-Thymine Watson-Crick base pair and Guanine-Cytosine Watson-Crick base pair are most stable in an aqueous environment.

**Key Words :** Watson-Crick DNA Base Pair, Quantum Mechanics, Integral Equation Theory of Liquids, Solvation Free Energy

### Introduction

Nucleic acids bases (Adenine, Cytosine, Guanine, and Thymine) exist as the main component of DNA double helix structures,<sup>1</sup> the essential molecules of life and evolution. Despite many kinds of nucleic base pairs are possible with different patterns of hydrogen bondings between the bases, the Adenine-Thymine and Guanine-Cytosine base pairs have been demonstrated as a promising component of DNA double helix structures.<sup>2</sup> Especially, Watson and Crick suggested that the A-T and G-C Watson-Crick base pairs are a specific component of DNA double helix structures.<sup>3</sup> The water effect of the nucleic base pairs are important in organism. However, the interactions between water molecule and nucleic base pairs have been elusive. The research that the Watson-Crick pairs are proper structures for DNA double helix structures in an aqueous phase is important.

To verify this idea, we optimized the four Adenine-Thymine base pairs (Hoogsteen pair, reverse Hoogsteen pair, Watson-Crick pair, and reverse Watson-Crick pair) and the three Guanine-Cytosine base pairs (GC 1, GC 2, and Watson-Crick pair) by using the density functional theory calculations at the B3LYP/6-31+G\*\* level.<sup>4-6</sup> Then, we studied Watson-Crick DNA base pairs of hydrolysis free energies solvated in water by using the 3D-RISM-SCF (three-dimensional reference interaction

site model self-consistent field) theory that integrates the ab initio quantum chemistry method and the integral equation theory of liquid.<sup>7-9</sup> The theory used the electronic structure of solute and the statistical solvent distribution around DNA base pairs in a self-consistent manner, and properly accounts for chemical specificities of solute-water interactions such as hydrogen bonding similarly to molecular simulations with explicit solvent. We quantitatively demonstrate the role of water in the hydrolysis of base pairs and we clarify nucleic base pairs-water interactions depending on the hydrolysis free energy.<sup>10-16</sup> We argue that why the experimental results for the aqueous-phase hydrolysis free energies are moderately that DNA double helix structures composed of Watson-Crick base pairs.

### Theory and Computational Method

The DFT (density-functional theory) at B3LYP/6-31+G\*\* level of the quantum chemistry theory were used for calculations of the geometry.<sup>13-16</sup> The interaction energy of the four Adenine-Thymine base pairs (Adenine-Thymine Hoogsteen base pair, Adenine-Thymine reverse Hoogsteen base pair, Adenine-Thymine Watson-Crick base pair, and Adenine-Thymine reverse Watson-Crick base pair) and three Guanine-Cytosine base pairs (Guanine-Cytosine 1 base

pair, Guanine-Cytosine 2 base pair, and Guanine-Cytosine Watson-Crick base pair) was optimized at the B3LYP/6-31+G\*\* level by using the Gaussian 03 program in a gas phase.<sup>4-6</sup>

Then, the mechanical calculation was used for stability structure in an aqueous phase. The 3D-RISM-SCF (three-dimensional reference interaction site model self-consistent field) was used for quantification the effect of water at 298.15 K and 1.0000 atm. The four Adenine-Thymine base pairs (Adenine-Thymine Hoogsteen base pair, Adenine-Thymine reverse Hoogsteen base pair, Adenine-Thymine Watson-Crick base pair, and Adenine-Thymine reverse Watson-Crick base pair) and three Guanine-Cytosine base pairs (Guanine-Cytosine1 base pair, Guanine-Cytosine2 base pair, and Guanine-Cytosine Watson-Crick base pair) was calculated by using equation of solvation free energy.<sup>7-9</sup>

In this method, the electronic reorganization energy  $E_{Reorg}$  is given by a subtraction of the aqueous phase free energy  $E_{UU}^{sol}$ , the gas phase free energy  $E_{UU}^{gas}$ :

$$E_{Reorg} = E_{UU}^{sol} - E_{UU}^{gas} \quad (1)$$

The total solvation free energy is shown by a sum of the solute energy  $E_{UU}$ , the solvation free energy (excess chemical potential)  $\Delta\mu$ :

$$F_{total \text{ solvation free energy}} = E_{UU} + \Delta\mu \quad (2)$$

$$E_{UU} = \langle \sum_{sol} \sum_{i,j} \frac{Z_i Z_j}{|r_i - r_j|} \rangle + \sum_{i < j}^{N_u} \frac{Z_i Z_j}{|r_i - r_j|} \quad (3)$$

$$\Delta\mu = \frac{\rho}{\beta} \sum_{j=1}^{N_v} \int dr \{ \sum c_j(r) - \frac{1}{2} h_j(r) c_j(r) +$$

$$\frac{1}{2} h_j(r)^2 \Theta(\sum h_j(r)) \} \quad (4)$$

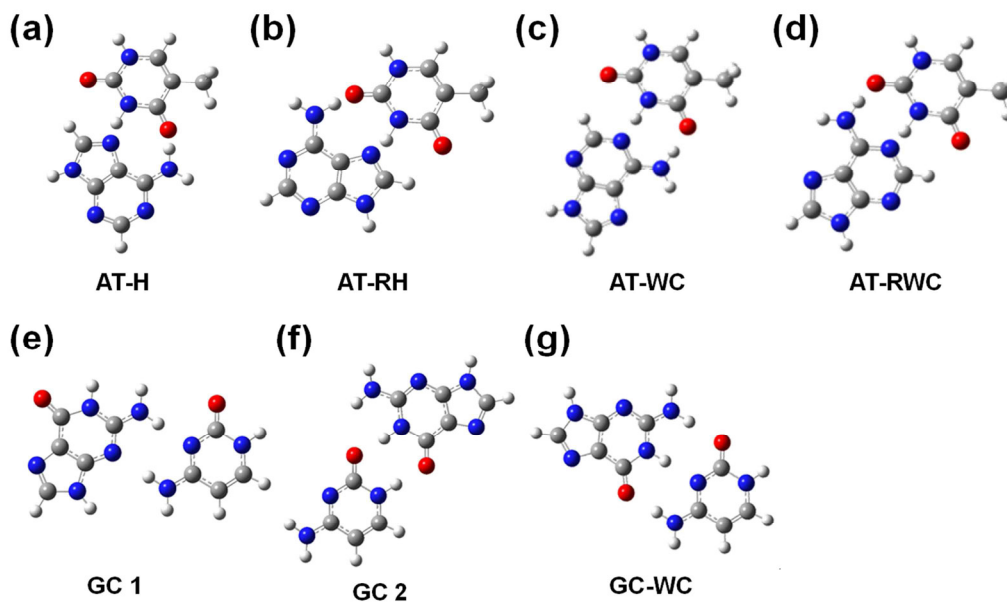
Here  $\Theta(\chi)$  is the Heaviside step function and  $h_j(r), c_j(r)$  refer to the 3D total and direct correlation functions of the water site  $j$  at position  $r$ .<sup>17-21</sup>

## Results and Discussion

### 1. Structural analysis

For the four possible states of Adenine-Thymine base pair and three possible states of Guanine-Cytosine base pair, the computed interaction energies in a gas phase at the B3LYP/6-31+G\*\* level are summarized and compared with experimental data in Table1.

The gas phase geometries of computation at the B3LYP/6-31+G\*\* level are shown in Figure1. As shown in Figure1, the four possible A-T base pairs (A-T H (a), A-T RH (b), A-T WC (c), A-T RWC (d)) have three hydrogen bindings each other. However, the only GC WC (e) has three hydrogen bindings and others (GC1 (f), GC2 (g)) have two hydrogen bindings each other. The three G-C base pairs have the planar structures. On the other hands, only the A-T WC base pair is the planar structure and the others are the bent structures. The results in the gas phase can be explained in stabilization of binding in base pairs. Compared to the gas-phase geometries, the computed structures for the A-T H base pair and G-C WC base pair show lower energy than other base pairs. The A-T and the GC WC base pair can composed of the DNA helix structures in a gas phase because these structures are composed with planar base pairs.



**Figure 1.** Optimized molecular geometries of (a) AT-H, (b) AT-RH, (c) AT-WC, (d) AT-RWC, (e) GC1, (f) GC2, and (g) GC-WC

## 2. Solvation Free Energy

The solvent effect of the base pairs based on the 3D-RISM-SCF theory is summarized in Table 2. The lowest solvation free energy among Adenine-Thymine base pairs is Reverse Hoogsteen base pair and among Guanine-Cytosine base pairs is Watson-Crick base pair and these structure more stable than others. The DNA helix structures also composed with planar G-C WC and A-T WC pairs that are different from bent A-T H,

A-T RH, and A-T RWC pairs in an aqueous phase.<sup>22</sup> Depicted in Figure 2 are the three-dimensional distributions of water oxygen (O) and hydrogen (H) around all the solute species related to the reactions. Figure 3 shows the radial distribution functions of O and H of solvent water around the base pairs and shows quite characteristic hydration patterns attributed to water-hydrogen around the oxygen atoms in the base pairs.

**Table 1.** Gas phase energy computed by quantum mechanics

	Absolute energy (Hartree)	Interaction energy (kcal/mol)	Relative energy (kcal/mol)
AT-H	-921.544996	-13.56	0
AT-RH	-921.544646	-13.34	0.22
AT-WC	-921.541589	-11.42	2.14
AT-RWC	-921.543152	-12.40	1.16
GC1	-937.57591	-13.78	12.85
GC2	-937.59174	-23.71	2.92
GC-WC	-937.65964	-26.64	0.00

**Table 2.** Aqueous-Phase energy in kcal/mol

	E(UU)	$\Delta\mu$	F <sub>total free energy</sub>	E <sub>reorg</sub>	Interaction Energy	Relative Energy
AT-H	-577937.12	3.58	-577933.53	8.11	-0.45	0.74
AT-R	-577936.76	2.49	-577934.27	8.22	-1.19	0
AT-WC	-577934.16	1.32	-577932.84	8.83	-0.25	1.43
AT-RWC	-577934.85	0.70	-577934.15	9.12	-1.06	0.12
GC1	-587984.95	-25.72	-588010.66	17.57	-0.002	-1.29
GC2	-587999.94	-12.54	-588012.48	12.24	-0.005	-3.11
GC-WC	-588004.23	-10.21	-588014.44	11.09	-0.008	-5.06

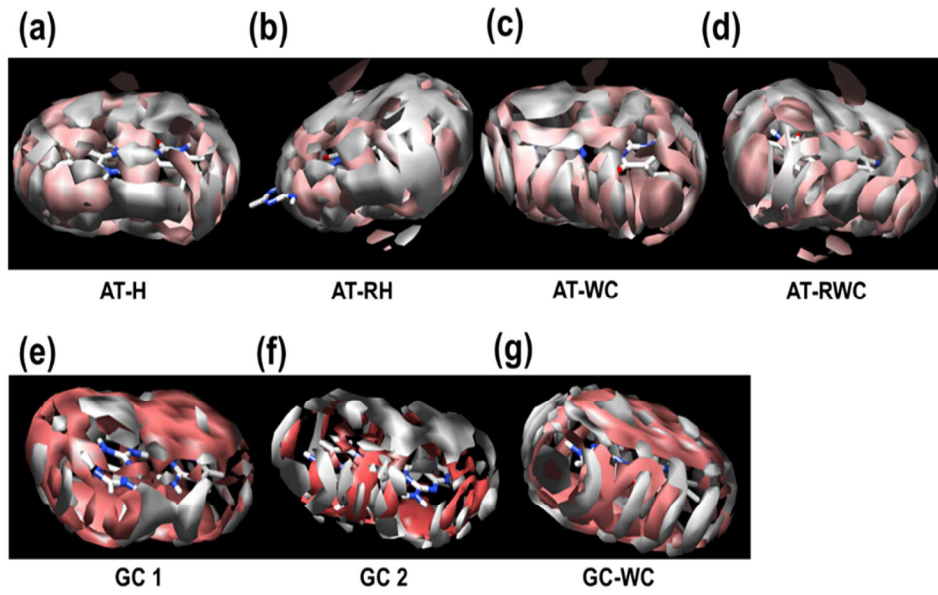


Figure2. Water distribution of (a) AT-H, (b) AT-RH, (c) AT-WC, (d) AT-RWC, (e) GC1, (f) GC2, and (g) GC-WC

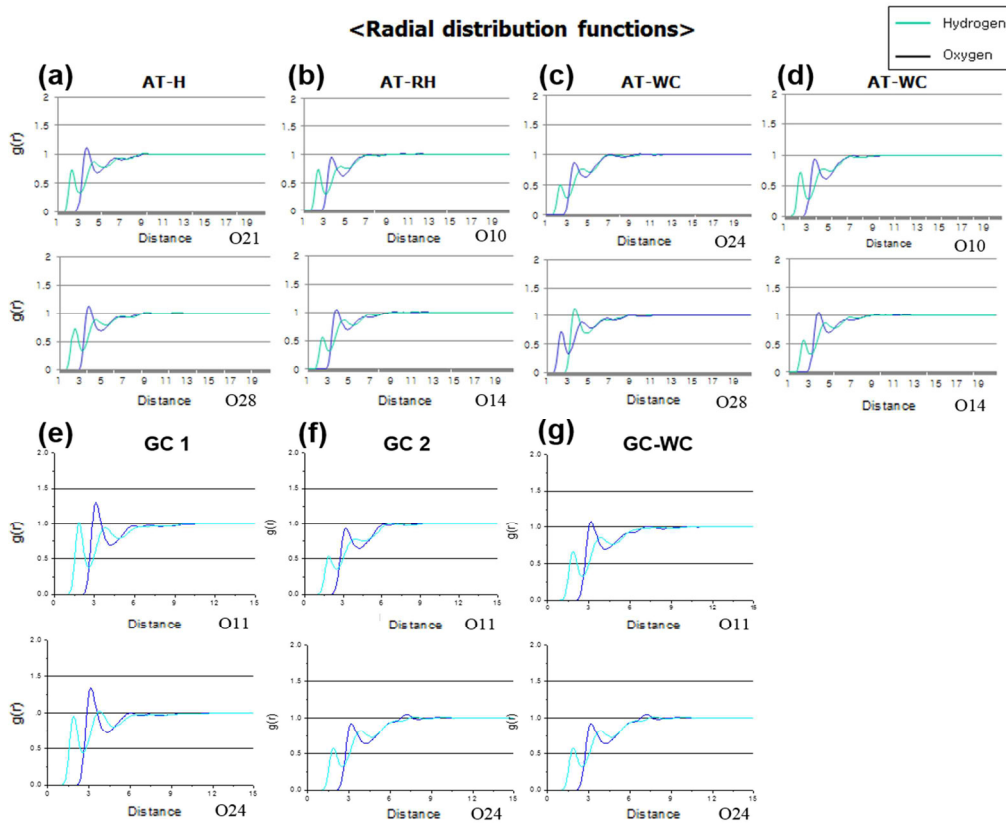


Figure3. Radial distribution functions of (a) AT-H, (b) AT-RH, (c) AT-WC, (d) AT-RWC, (e) GC1, (f) GC2, and (g) GC-WC

### Conclusion

The structures of four Adenine-Thymine base pairs and three Guanine-Cytosine base pairs are computationally studied by using density functional theory at the B3LYP/6-31+G\*\* level as well as by 3D-RISM (three-dimensional reference interaction site model) theory. The Adenine-Thymine Hoogsteen base pair is the most stable structure than other three base pairs (Adenine-Thymine reverse Hoogsteen base pair, Adenine-Thymine Watson-Crick base pair, and Adenine-Thymine reverse Watson-Crick base pair) in a gas phase. On the other hand, the Adenine-Thymine reverse Hoogsteen base pair is most stable structure than other three base pairs (Adenine-Thymine Hoogsteen base pair, Adenine-Thymine Watson-Crick base pair, and Adenine-Thymine reverse Watson-Crick base pair) in an aqueous phase. Then, The Guanine-Cytosine Watson-Crick base pair (GCWC) is more stable than the other two pairs (GC1 and GC2) not only in a gas phase but also in an aqueous phase. Based on the solvation free energy analysis, GCWC compared to GC1 and GC2 would be proper to make a DNA helix structure.

### Acknowledgements

This research was supported by the EDISON Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Science, ICT & Future Planning (No. NRF-2012-M3C1A6035357).

### References

- Sýponer, J.; Florián, J.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn.* 1996, 13, 827-833.
- Watson, J. D.; Crick, F. H. C. *Cold Spring Harbor Symposia Quant. Biol.*, 1953, 18, 123
- Watson, J. D.; Crick, F. H. C. *Nature* 1953, 171, 737-738
- Becke, A. D. *J. Chem. Phys.* 1993, 98, 5648-5652.
- 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.; Burant, J. C.; Millam, J. M.; Iyengay, 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.; Adamo, C.; Jaramillo, J.; Comperts, R.; Startmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenbuerg, 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.; Pople, J. A. *Gaussian 03, Revision C.02*; Gaussian, Inc.: Wallingford, CT, 2004.
- Kovalenko, A.; Hirata, F. *J. Chem. Phys.* 1999, 110, 10095-10112.
- Sato, H.; Kovalenko, A.; Hirata, F. *J. Chem. Phys.* 2000, 112, 9463-9468
- Yoshida, N.; Hirata, F. *J. Mol. Liq.* 2011, 159, 83-92
- Kohn, W.; Sham, L. *J. Phys. Rev.* 1965, 140, A1133-A1138.
- George, P.; Witonsky, R. J.; Trachtman, M.; Wu, C.; Dorwart, W.; Richman, L.; Richman, W.; Shurayh, F.; Lentz, B. *Biochim. Biophys. Acta*, 1970, 223, 1-15.
- Hofmann, K. P.; Zundel, G. *Experimentia* 1974, 30, 139-140
- de Meis, L. *Biochim. Biophys. Acta* 1989, 973, 333-349.
- Romero, P. J.; de Meis, L. *J. Biol. Chem.* 1989, 264, 7869-7873.
- Saint-Martin, H.; Ortega-Blake, I.; Les, A.; Adamowicz, L. *Biochim. Biophys. Acta* 1994, 1207, 12-23.
- Colvin, M. E.; Evleth, E.; Akacem, Y. *J. Am. Chem. Soc.* 1995, 117, 4357-4362.
- Akola, J.; Jone, R. O. *J. Phys. Chem. B* 2003, 107, 11774-11783
- Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* 1993, 97, 10269-10280.
- Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* 1987, 91, 6269-6271.
- Pettitt, B. M.; Rossky, P. J. *J. Chem. Phys.* 1982, 77, 1451-1457.
- Jooyeon Hong, Norio Yoshida, Song-Ho Chong, Chewook Lee, Sihyun Ham, and Fumio Hirata., *J. Chem. Theory Comput.* 2012, 8, 2239-2246
- Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* 2009, 113, 6387-6396.
- Maidier Machado, Pablo Ordejón, Emilio Artacho, Daniel Sánchez-Portal and José M. Soler, *Physics.chem-ph*, 1999, 11.