Computer-Aided Molecular Design of Hydrogen Bond Equivalents of Nucleobases: Theoretical Study of Substituent Effects on the Hydrogen Bond Energies of Nucleobase Pairs

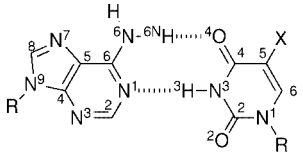
Shun-ichi Kawahara[a] and Tadafumi Uchimaru*[a]

Keywords: Ab initio calculations / DNA / Gene technology / Hydrogen bonds / RNA / Substituent effects

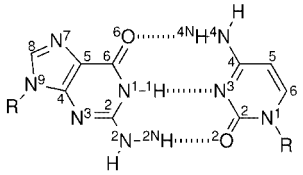
Substituent effects on the hydrogen bond energies of Watson–Crick-type base pairs, formed between a chemically modified nucleic acid base derivative and an unmodified one, were evaluated by ab initio molecular orbital theory. Different trends were observed in the relationship between the substituent and the hydrogen bond energy in each base pair. The predicted hydrogen bond energies correlated well with the experimentally measured binding properties, and so ab initio calculation appears to be an effective method with which to estimate the stabilities of base pairs between chemically modified nucleic acid bases.

Introduction

Hydrogen bond formation in a Watson−Crick-type base pair[1] (Figure 1) is essential for transmission of genetic information (i.e., the processes of transcription from DNA to mRNA,[2] and of translation from mRNA to protein via tRNA, based on highly selective molecular recognition).[3]



A-U (X=H), A-T (X=Me)



G-C

Figure 1. Watson−Crick base pairs

It has widely attracted much attention: for the construction of artificial supermolecular systems,[4] for template synthesis,[5] and for antisense technology,[6] for example. A molecule capable of selectively forming a stable complex is needed for these applications. Modification of base moieties has to be taken into account in the molecular design,[7] but no systematic study, except for our theoretical investigations,[8] has yet been made in pursuit of improvement of the base pair stability.

Theoretical studies are helpful for understanding the nature of the hydrogen bond and for prediction of the hydrogen bond stability of a modified base pair. We have carried out computer-aided molecular design of nucleic acid base analogues with the goal of controlling the base pair stability. In this account we describe an ab initio study regarding the substituent effect on hydrogen bond energies in base pairs composed of one chemically modified nucleic acid base derivative and one unmodified one.

Results

Computational Method for the Estimation of Hydrogen Bond Energies

Previous theoretical studies[9] on the hydrogen bond energies of the natural nucleic acid base pairs have been reviewed by Sponer and Hobza.[9a]

In most theoretical studies, the hydrogen bond energies of the Watson−Crick type base pairs were evaluated at the second-order Møller-Plesset (MP2) level of theory by the use of double-ζ basis sets with polarization.[9] The contribution of higher-level electron correlation to the calculated values of hydrogen bond energies was small.[10] The values of the hydrogen bond energies of A−U and G−C base pairs, evaluated at the computational levels of MP2/6-31+G(2d′,p′)[11]//HF/6-31G(d,p)[12] (−13.16 kcal/mol for A−U and −26.08 kcal/mol G−C), which we adopted in this study, were in excellent agreement with values calcu-

[a] National Institute of Advanced Industrial Science and Technology (AIST), AIST Tsukuba Central, Tsukuba, 305-8565, Japan
Fax: (internat.) + 81-29/861-4487
E-mail: t-uchimaru@aist.go.jp

lated at MP2/6-311++G(3df,p)//HF/6-311++G(3d,p) (−13.48 kcal/mol for A−U and −26.58 kcal/mol G−C).[13] The errors in calculated values for hydrogen bond energies, which originate from incompleteness of the basis set, should be comparable for all base pairs in the case of the same kind of base pair.

The hydrogen bond energies of the base pairs were evaluated by a supermolecular method. The basis set superposition error (BSSE) was corrected for by use of the counterpoise method.[14] Hereafter we refer to the molecular interaction energy without and with BSSE correction as δE and as ΔE^{HB}, respectively [Equations (1) and (2)]. A more negative ΔE^{HB} value means a more stable hydrogen bond. The term $\Delta\Delta E$ was defined as the substituent effect on ΔE^{HB} [Equation (3)].

$$\delta E(\text{M-N}) = E(\text{M-N}) - \{E(\text{M}) + E(\text{N})\} \quad (1)$$

$$\Delta E^{HB}(\text{M-N}) = \Delta E(\text{M-N}) + \text{BSSE} \quad (2)$$

$$\Delta\Delta E = \Delta E^{HB}(\text{M}-\text{N}) - \Delta E^{HB}(\text{A}-\text{U or G}-\text{C}) \quad (3)$$

M: Modified base; A^X, U^X, G^X, and C^X.
N: Natural Watson−Crick complementary base; U, A, C, and G, respectively.
M−N: A^X−U, U^X−A, G^X−C, and C^X−G.

C_s symmetry was assumed in structural optimization. Non-planarity of the exocyclic amino moiety in the isolated bases has been reported,[9a] but the differences in the energy derived from the planar structures and from the non-planar structures [ΔE^{NP}, Equation (4)] of A and C were very small and the structures of the bases in the Watson−Crick type base pairs were planar.[9a] The change in hydrogen bond energy originating from non-planarity of the bases is thus negligible, except in the case of the G^X−C base pairs. For the G^X−C base pairs, we considered ΔE^{NP}. The ΔE^{NP} values were calculated at the MP2/6-31+G(2d′,p′)//MP2/6-31G(d,p) level of calculation.

$$\Delta E^{NP} = E(G^X_{Planar}) - E(G^X_{Nonplanar}) \quad (4)$$

Conformational isomers were produced when the substituents were introduced on the exocyclic amino moiety (Figure 2). The conformers **I** can form a Watson−Crick-type base pair, while the conformers **II** cannot. The value of ΔE^{Rot}(M) was determined by use of Equation (5) for such derivatives. A positive ΔE^{Rot} value shows that the conformer **I** is unstable.

$$\Delta E^{Rot}(\text{M}) = E(\text{M}^{\text{I}}) - \text{E}(\text{M}^{\text{II}}) \quad (5)$$

ΔE^{Total}, which is the molecular interaction energy including the exocyclic amino moiety rotation and the energy difference originating from the non-planarity, is calculated as in Equation (6).

$$\Delta E^{Total} = \Delta E^{HB}(\text{M}-\text{N}) + \Delta E^{Rot}(\text{M}-\text{N}) + \Delta E^{NP} \quad (6)$$

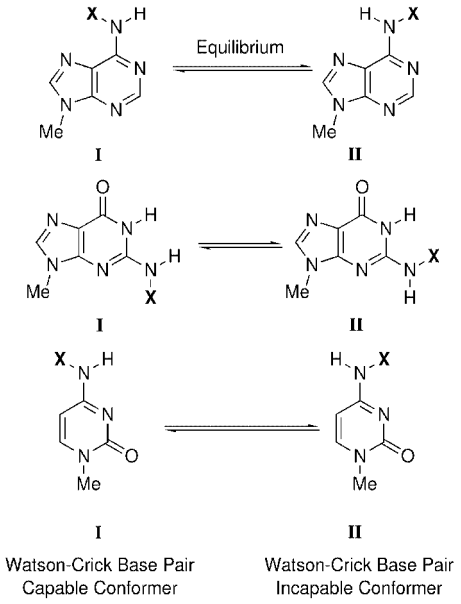


Figure 2. Exocyclic bond rotation in A^X, G^X, and C^X

The calculations were carried out with the aid of the SPARTAN[15] and the GAUSSIAN 94 programs.[16]

Classification of the Base Modification

The atoms are labeled as in Figure 1. The type of base modification is classified into the following six groups:

Group A: Unmodified base.

Group B: An electron-withdrawing group (EWG) or an electron-donating group (EDG) was introduced on the endocyclic carbon atom of the purine or pyrimidine.

Group C: An EWG or an EDG was introduced on the exocyclic amino moiety on the base.

Group D: The exocyclic oxygen atom is replaced by a sulfur atom, the exocyclic amino moiety is deleted and/or a new exocyclic amino moiety is added. The hydrogen bond of the exocyclic moiety should be weakened, removed, or added in this group.

Group E: A carbon or a nitrogen atom on the heterocycles is replaced with another atom.

Group F: Bases possessing larger conjugate systems. Detailed classifications of the derivatives are shown in Figures 4, 7, 9, and 11.

Substituent Effects on A^X−U Base Pairs

There are two hydrogen bonds between A^X and U (Figure 3). A^X acts as an electron acceptor in H-bond A and as an electron donor in H-bond B. A decrease in the electron population of the purine ring thus strengthens the H-bond A and weakens the H-bond B, and vice versa. Namely, the substituent effects on the strengths of H-bonds A and B may cancel each other out, so the substituent effect on the total hydrogen bond energy of A^X−U is difficult to forecast

from the structure. Classification and abbreviations of A^Xs are shown in Figure 4.[17]

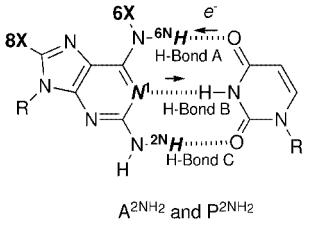


Figure 3. Hydrogen bond between A^X and U

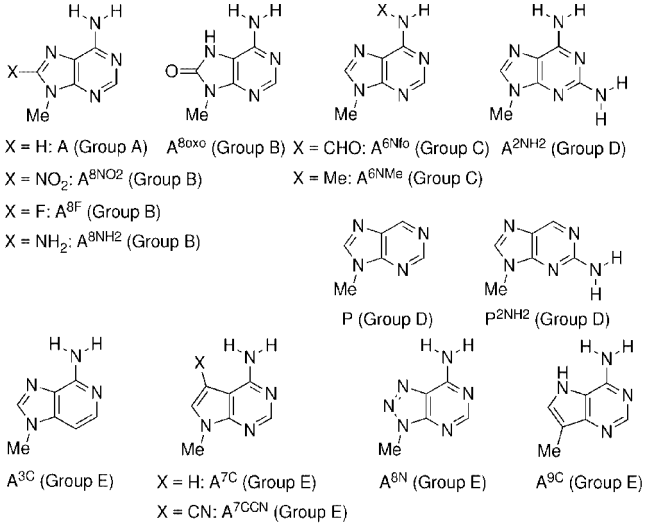


Figure 4. Substituent-introduced 9-methyl adenine derivatives [A^X]

Table 1 shows the results of the theoretically estimated hydrogen bond properties of each A^X−U base pair. No significant trend was observed in the relationship between the substituent in adenine derivatives and the hydrogen bond energies. Both EWGs (A^{8NO2}, A^{8oxo}, and A^{8F}) and EDGs (A^{8NH2}) on the 8-position of A stabilize the hydrogen bond with U (Figure 5). The presence of a methyl group on the exocyclic amino group in the 6-position (A^{6NMe}) has little effect on the hydrogen bond stability. On the other hand, the presence of a formyl group at the same position (A^{6Nfo}) destabilizes the hydrogen bond. For both A^{6NMe} and A^{6Nfo} the conformer **I** was found to be higher in energy (unstable) than the conformer **II**. $\Delta E^{HB}(P^{2NH2}-U)$ is less negative than $\Delta E^{HB}(A-U)$. As would be expected, $\Delta E^{HB}(A^{2NH2}-U)$, which forms three hydrogen bonds, is more negative than $\Delta E^{HB}(A-U)$. $\Delta E^{HB}(P-U)$, which forms only one hydrogen bond, is less negative than $\Delta E^{HB}(A-U)$. The substituent effects on ΔE^{HB} in Group E were not so large.

Substituent Effects on U^X−A Base Pairs

With regard to the hydrogen bond structure (Figure 6), the effect of the electron population of the pyrimidine ring of U should be opposite to that of the purine ring of A. As in the case of A^X, the substituent effects on the strengths

Table 1. ΔE^{HB}, BSSE, $\Delta\Delta E$, ΔE^{Rot}, and ΔE^{Total} [kcal/mol] of each A^X, and the binding constants k [l/mol]

Group	A^X	ΔE^{HB}	BSSE	$\Delta\Delta E$	ΔE^{Rot}	ΔE^{Total}	k[a]
A	A	−13.11	2.13	−			100
	A^{8NO2}	−15.37	2.21	−2.26			
B	A^{8oxo}	−13.16	2.16	−0.05			
	A^{8F}	−13.20	2.17	−0.09			140[b]
	A^{8NH2}	−14.25	2.16	−1.14			
C	A^{6Nfo}	−10.87	2.33	+2.24	0.08	−10.79	
	A^{6NMe}	−12.83	2.30	+0.28	1.83	−11.00	50
	P	−8.42	1.73	+4.69			
D	P^{2NH2}	−11.96	2.19	+1.15			45
	A^{2NH2}	−14.96	2.50	−1.85			170
	A^{3C}	−12.31	2.20	+0.80			
	A^{7C}	−13.43	2.19	−0.32			
E	A^{7CCN}	−12.98	2.21	+0.13			
	A^{8N}	−13.30	2.19	−0.19			
	A^{9C}	−12.93	2.16	+0.18			

[a] Ref.[19] [b] The value of k for A^{8Br}.

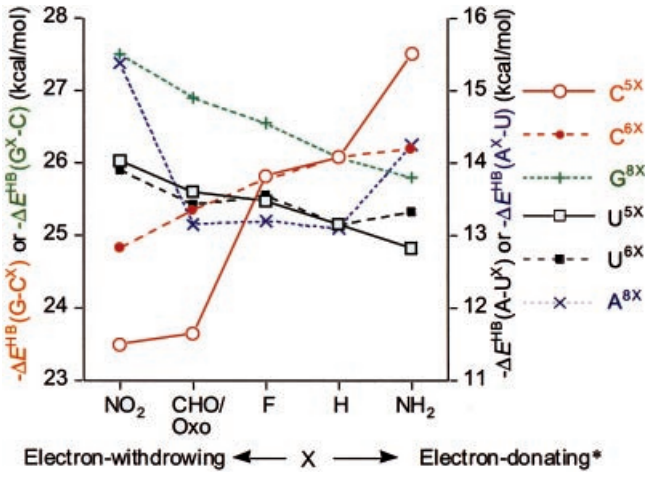


Figure 5. Substituent effect on ΔE^{HB} in groups B of the base pairs (* ref.[18])

of H-bonds A and B may cancel each other out. Figure 7 shows the classification and abbreviations for the uracil analogues.

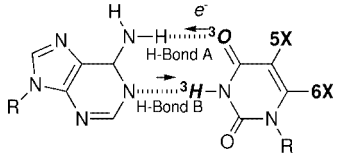


Figure 6. Hydrogen bond between U^X and A

Table 2 shows the results for the theoretically estimated hydrogen bond properties of each U^X−A base pair. In contrast to the substituent effect on hydrogen bond energy in A, there was a remarkable tendency for U^X species possessing more strongly EWGs to form the more stable base pairs with A (Figure 5). In view of the fact that U^X acts as an electron donor in H-bond A and as an electron acceptor

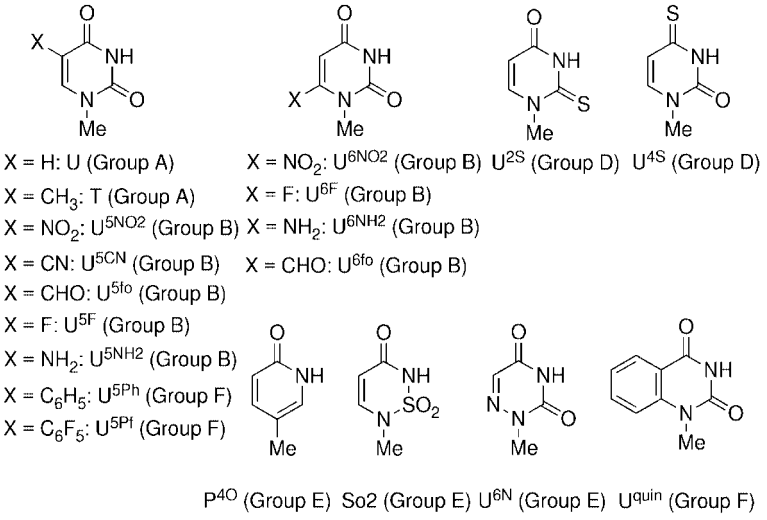


Figure 7. Substituent-introduced 1-methyl uracil derivatives [U^X]

in H-bond B, it is deduced that the H-bond B plays a more important role than H-bond A in these systems.

Table 2. ΔE^{HB}, BSSE, and $\Delta\Delta E$ [kcal/mol] of each U^X

Group	U^X	ΔE^{HB}	BSSE	$\Delta\Delta E$
A	U	−13.11	2.13	−
	T	−13.29	2.21	−0.18
	U^{5NO2}	−14.03	2.35	−0.92
	U^{5fo}	−13.60	2.25	−0.49
	U^{5F}	−13.47	2.23	−0.36
B	U^{5NH2}	−12.83	2.21	+0.28
	U^{6NO2}	−13.90	2.30	−0.79
	U^{6fo}	−13.42	2.30	−0.31
	U^{6F}	−13.55	2.23	−0.44
	U^{6NH2}	−13.33	2.22	−0.22
D	U^{2S}	−12.92	2.27	+0.19
	U^{4S}	−11.59	2.24	+1.52
	U^{6N}	−13.80	2.19	−0.69
E	P^{4O}	−14.77	2.00	−1.66
	So2	−14.61	2.74	−1.50
	U^{quin}	−13.20	2.28	−0.09
F	U^{5Ph}	−13.92	2.38	−0.81
	U^{5Pf}	−14.50	2.48	−1.39

The effect of replacement of the 4-position exocyclic oxygen atom by a sulfur atom was large, as expected. The effect of replacing the 2-position exocyclic oxygen atom by a sulfur atom was small.

The ΔE^{HB} value for U^{6N} was more negative than that for U. Replacement of a carbon atom with a nitrogen causes the heterocyclic ring to become electron-deficient, so the change in the ΔE^{HB} value for U^{6N} was in agreement with the substituent effect of ΔE^{HB} in Group B. For the same reason, ΔE^{HB} of So2 (Group E) was more negative than that of U. On the other hand, P^{4O}, which is regarded as a more electron-rich heterocyclic ring than U, formed a more stable base pair with A. The reason for the result for the P^{4O}−A base pair can be described on the basis of the secondary interaction proposal.[19]

The ΔE^{HB} value for U^X in Group F was more negative than that of U, while $\Delta\Delta E$ of U^{quin} and U^{5Ph} were very different. It is difficult to explain the difference in the $\Delta\Delta E$ values of U^{quin} and U^{5Ph}. The pentafluorophenyl group in U^{5Pf} acts as a strongly EWG. In U^{5Ph} and U^{5Pf}, large dihedral angles (46° and 61°, respectively) between the rings were observed, so the steric effect on the backbone structure of DNA and RNA should be serious.

Substituent Effects on G^X−C Base Pairs

Guanine derivatives act as electron-donors in H-bond A, but on the other hand, they act as electron-acceptors in H-bonds B and C, as shown in Figure 8, so it is considered that introduction of an EDG on G^X should make H-bond A stronger and, conversely, H-bonds B and C weaker, and vice versa. G^Xs were classified as shown in Figure 9. Table 3 shows the results of the theoretically estimated hydrogen bond properties of each G^X.

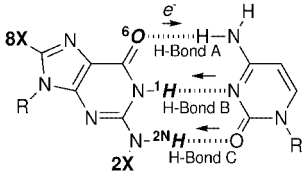


Figure 8. Hydrogen bond between G^X and C

A remarkable trend in ΔE^{HB} was observed: the substituent effect of the introduction of an EWG on the 8-position and on the exocyclic amino moiety strengthened the hydrogen bond. Thus, the sum of the substituent effects on the H-bonds B and C overcomes the substituent effect on the H-bond A. Consequently, $\Delta E^{HB}(G^{8NO2}-C)$ was the most negative in this study, and $\Delta E^{HB}(G^{8NH2}-C)$ was the least negative in group B. This trend was, as expected, the opposite of the substituent effect in cytosine (see Figure 5). The fluctuation in ΔE^{HB} resulting from the introduction of the substituent in G^{8X} was smaller than that in C^{5X}. Be-

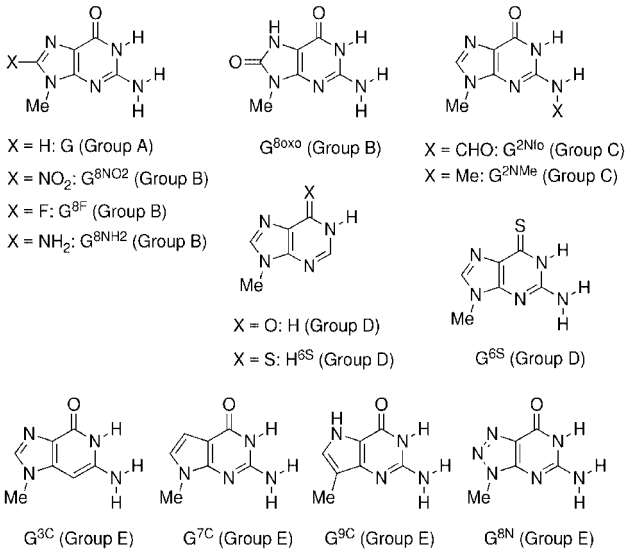


Figure 9. Substituent-introduced 9-methyl guanine derivatives [G^X]

Table 3. ΔE^{HB}, BSSE, $\Delta\Delta E$, ΔE^{Rot}, ΔE^{NP}, and ΔE^{Total} [kcal/mol] of each G^X

Group	G^X	ΔE^{HB}	BSSE	$\Delta\Delta E$	ΔE^{Rot}	ΔE^{NP}	ΔE^{Total}
A	G	−26.08	2.83	−		1.47	−24.61
	G^{8NO2}	−27.49	2.91	−1.41		2.88	−24.62
B	G^{8oxo}	−26.90	2.89	−0.82		1.37	−25.53
	G^{8F}	−26.54	2.88	−0.46		1.38	−25.16
	G^{8NH2}	−25.79	2.86	+0.29		5.78	−20.00
C	G^{2Nfo}	−26.68	3.18	−0.60	−0.99	0.00	−26.68
	G^{2NMe}	−26.06	3.01	+0.02	−1.62	0.00	−26.05
	G^{6S}	−23.52	2.66	+2.56		1.36	−22.16
D	H	−20.25	2.40	+5.83		−	−20.25
	H^{6S}	−19.24	2.45	+6.84		−	−19.24
	G^{3C}	−25.75	2.82	+0.33		2.55	−23.21
E	G^{7C}	−25.19	2.83	+0.89		1.77	−23.43
	G^{9C}	−24.19	2.79	+1.89		2.08	−22.11
	G^{8N}	−26.73	2.84	−0.65		1.11	−25.62

cause of the large $\Delta E^{NP}(G^{8NO2}-C)$ value, $\Delta E^{Total}(G^{8NO2}-C)$ was almost the same as that of the G−C base pair. The substituent effect on the exocyclic amino moiety of the guanine derivatives was also the opposite of that of the corresponding cytosine derivatives. For both G^{2NMe} and G^{2Nfo}, the conformer **I** was found to be energetically more stable than the conformer **II**. For discussion of a typical substituent effect on hydrogen bond energy, we mainly discuss ΔE^{HB}. The exocyclic amino moiety on G^{2Nfo} and G^{2NMe} was planar. Both $\Delta E^{HB}(H-C)$ and $\Delta E^{HB}(G^{6S}-C)$ were less negative than those of the G−C base pair. The effect of the deletion of the exocyclic amino moiety was larger than the effect of replacement of the exocyclic oxygen by sulfur. Although H−C forms only two hydrogen bonds, $\Delta E^{HB}(H-C)$ was more negative than $\Delta E^{HB}(A-U)$, while $\Delta E^{HB}(H^{6S}-C)$ was less negative than $\Delta E^{HB}(H-C)$. In Group E, only ΔE^{HB} of the $G^{8N}-C$ base pair was more negative than that of the G-C. The substituent from aromatic carbon to nitrogen causes the aromatic ring to become electron-deficient, so it is equivalent to introducing an EWG on the aromatic ring.

Substituent Effects on C^X-G Base Pairs

Cytosine derivatives act as electron-acceptors in H-bond A and as electron-donors in H-bond B and C, as shown in Figure 10, so it is considered that introduction of an EWG on C^X should make H-bond A stronger and, conversely, H-bonds B and C weaker, and vice versa. Thus, if the contributions of each of the three hydrogen bonds to the stability of the base pair are the same, the stability of the base pair should be strengthened by the introduction of an EDG. C^Xs were classified as shown in Figure 11.

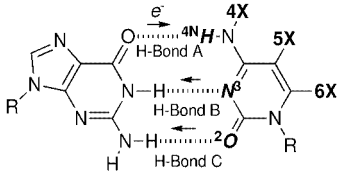


Figure 10. Hydrogen bond between C^X and G

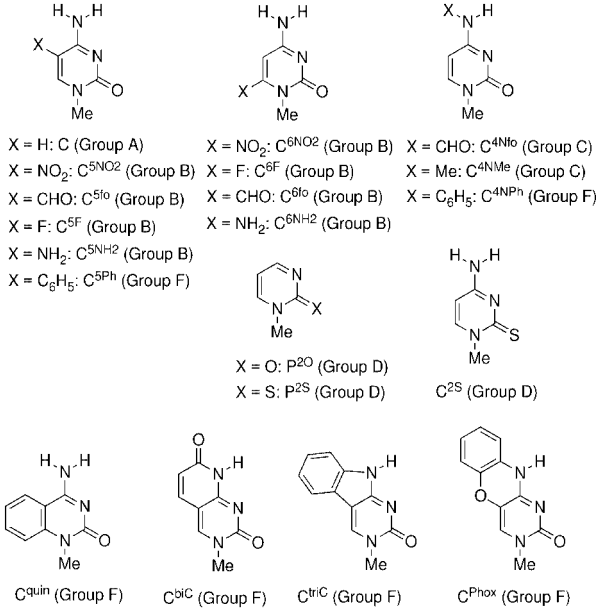


Figure 11. Substituent-introduced 1-methyl guanine derivatives [C^X]

Table 4 shows the results of the hydrogen bond properties of each C^X, while Figure 5 shows the substituent effects on ΔE^{HB} in groups B of C^X-G and the corresponding U^X-A. A remarkable tendency was observed: unlike the substituent effect in uracil, a C^X species possessing a more strongly EWG forms a less stable base pair with G. This result shows that the sum of the substituent effects on the H-bonds B and C overcomes the substituent effect on the H-bond A. The substituent effect on the 5-position was larger than that on the 6-position or that of the corresponding U^{5X} (Fig-

ure 5). It is considered that the formation of the intramolecular hydrogen bond enhances the electron-withdrawing ability of the substituent by accepting the hydrogen bond (Figure 12).

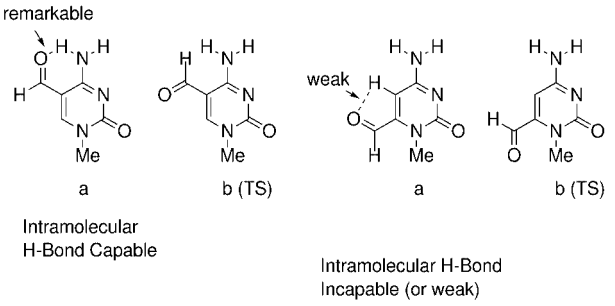


Figure 12. Intramolecular hydrogen bond formation in C^{5fo} and C^{6fo}

Table 4. ΔE^{HB}, BSSE, $\Delta\Delta E$, ΔE^{Rot}, ΔE^{NP}, and ΔE^{Total} [kcal/mol] of each C^X

Group	C^X	ΔE^{HB}	BSSE	$\Delta\Delta E$	ΔE^{Rot}	ΔE^{Total}
A	C	−26.08	2.83	−		
	C^{5NO2}	−23.49	2.91	+2.59		
	C^{5NO2} [a]	−23.41	2.91	+2.67		
	C^{5For}[a][b]	−23.64	2.88	+2.44		
	C^{5For}[b][a]	−24.85	2.89	+1.23		
	C^{5F}	−25.82	2.90	+0.26		
B	C^{5NH2} [a]	−27.51	2.94	−1.43		
	C^{5NH2}	−25.84	2.92	+0.24		
	C^{6NO2}	−24.84	2.91	+1.24		
	C^{6For}[a]	−25.36	2.87	+0.72		
	C^{6For}[b][a]	−25.73	2.88	+0.35		
	C^{6F}	−25.79	2.43	+0.29		
	C^{6NH2}	−26.21	2.89	−0.13		
C	C^{4Nfo}	−26.05	3.12	+0.03	−1.43	
	C^{4NMe}	−27.09	3.04	−1.02	1.12	−25.97
	C^{2S}	−23.58	2.57	+2.50		
D	P^{2O}	−17.77	2.26	+8.31		
	P^{2S}	−18.10	2.34	+7.98		
	C^{4NPh}	−27.23	3.37	−1.15	−0.16	
	C^{5Ph}	−25.75	3.07	+0.33		
	C^{triC}	−27.23	3.18	−1.15		
F	C^{quin} [a]	−26.31	2.99	−0.23		
	C^{quin}	−26.30	2.99	−0.22		
	C^{biC}	−21.51	3.06	+4.57		
	C^{Phox}	−28.35	3.42	−2.27		

[a] *C*s structure of C^{5NO2}, C^{5For}[b], C^{5NH2}, C^{6For}[b], and C^{quin} was not the energy minimum. [b] See also Figure 12 for the difference of $C^{5/6For}$[a] and $C^{5/6For}$[b].

Alkylation of the exocyclic amino group on the 4-position (^4N) strengthened the hydrogen bonding of the base pairs. Acylation on the same position had almost no effect.

Although the substituent effect in group F was relatively large, there was no remarkable trend in the substituent effect on ΔE^{HB} in group F. Except for that of C^{5Ph}, the ΔE^{HB} value of the group E was more negative than that of the C−G base pair. ΔE^{HB} of C^{Phox}−G was the most negative in this study.

As would be expected, the replacement of exocyclic oxygen with sulfur results in a decrease of hydrogen bond energy of the base pair. This effect was larger for C^{2S} than for U^{4S}. The value of $\Delta\Delta E$ (G−C^{2S}) was almost the same as $\Delta\Delta E$ (G^{6S}−C). The effect of the deletion of the exocyclic amino moiety of C (G−C → G−P^{2O}, Figure 11 and Table 4) was larger than that for G (G−C → H−C, Figure 9 and Table 3). According to the secondary interaction proposal,[19] the ΔE^{HB} value of the G−P^{2O} base pair was nevertheless more negative than that of the A−U base pair. Obviously, the hydrogen bond energy per hydrogen bond is different for each base pair, and the base pair stability is determined not only by the number and the direction of the hydrogen bonds in the base pair, but also by the hydrogen bond capability of the hydrogen bond sites.

Comparison Between the Predicted Hydrogen Bond Energies and Experimentally Observed Base Pair Stability

By examining changes in the N−H stretching mode in IR spectra, Kyogoku et al. derived the binding constants (k) between uracil and some adenine derivatives.[20] ΔE^{HB} was compared to the log of the binding constants (k) between uracil and some adenine derivatives (A, A^{2NH2}, P^{2NH2}, A^{8Br},[21] and A^{6NMe}). As shown in Figure 13, the theoretically estimated substituent effect on ΔE^{HB} reproduced the substituent effect on experimentally observed k: $A^{2NH2} > A^{8Br}/A^{8F} \approx A > A^{6NMe} \approx P^{2NH2}$.[22]

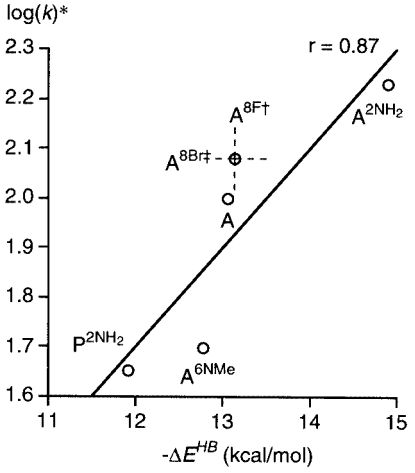


Figure 13. Relationship between theoretically estimated ΔE^{HB} and log of binding constant of some A^X species (* ref.[20] and † ref.[21])

Ishikawa et al. reported the correlation of the chemical shift of the imino proton, the pK_a value of the imino proton, the nitrogen-proton [N^1(A)-^{3}H(U)] coupling constant through the covalent bond ($^1J_{NH}$), and ΔE^{HB}.[23] These factors correlate with each other, especially between ΔE^{HB} and $^1J_{NH}$, and between the chemical shift and the pK_a value of the imino proton.

The duplex stabilities of the oligonucleotides were observed as melting temperatures (T_ms). An increase in T_m

shows an increase in the duplex stability of M introduced oligonucleotides, and vice versa. A change in the T_m is therefore an index of the substituent effect on duplex stability. However, T_ms and changes in T_ms are highly sensitive to the experimental conditions, and the experimental conditions of these studies differed. Comparison between $\Delta\Delta E$ and change in T_m is therefore not discussed here.

There have been no systematic studies investigating the experimental binding properties (e.g., values of $^1J_{NH}$[23]) and binding constants (k)[20] of guanine derivatives and cytosine derivatives, so systematic discussion of the relationship between theoretical prediction and experimentally ascertained binding properties for guanine and cytosine derivatives is impossible.

Conclusion and Outlook

We have described substituent effects on the hydrogen bond energies of Watson−Crick-type base pairs formed between a chemically modified nucleic acid base derivative and an unmodified one. The presence of an electron-withdrawing or an electron-donating group may effectively change the stability of the base pair, and the theoretically estimated stability was in agreement with experimental results in most cases. The results show how the stability of the base pair may be strengthened or weakened through the introduction of substituents. We have demonstrated a methodology to control the hydrogen bond stability of a nucleic acid base by use of computer-aided molecular design. In the last decade, new types of base pairs, using non-natural bases, have been developed by some groups.[24] Only the hydrogen bond equivalents of natural nucleic acid bases are described in this account, but they should be usable for the molecular design of unnatural base pairs.[25]

Acknowledgments

S.-i. K. and T. U. are indebted to Prof. M. Sekine, Prof. K. Taira, Prof. T. Wada, Dr. S. Kawauchi, and Dr. A. Kobori for their valuable contributions to this research. Mr. Y. Takagi, Mr. Y. Ikeda, and Miss M. Shiota are acknowledged for their helpful discussions.

[1] J. D. Watson, H. C. Crick, *Nature* **1953**, *171*, 737−738.
[2] S. Brenner, F. Jacob, M. Meselson, *Nature* **1961**, *190*, 576−581.
[3] J. J. Hopfield, *Proc. Natl. Acad. Sci. U. S. A.* **1974**, *71*, 4135−7139.
[4] [4a] J. L. Sessler, R. Wang, *Angew. Chem. Int. Ed.* **1998**, *37*, 1726−1729; *Angew. Chem.* **1998**, *110*, 1818−1821. [4b] J. J. González, P. Prados, J. de Mendoza, *Angew. Chem. Int. Ed.* **1999**, *38*, 525−528; *Angew. Chem.* **1999**, *111*, 546−549.
[5] [5a] T. Tjivikua, P. Ballester, J. Rebek, Jr., *J. Am. Chem. Soc.* **1990**, *112*, 1249−1250. [5b] J. S. Nowick, Q. Feng, T. Tjivikua, P. Ballester, J. Rebek, Jr., *J. Am. Chem. Soc.* **1991**, *113*, 8831−8839. [5c] F. M. Menger, A. V. Eliseev, N. A. Khanjin, M. J. Sherrod, *J. Org. Chem.* **1995**, *60*, 2870−2878. [5d] L. J. Prins, D. N. Reinhoudt, P. Timmerman, *Angew. Chem. Int. Ed.* **2001**, *40*, 2382−2426; *Angew. Chem.* **2001**, *113*, 2446−2492.
[6] [6a] J. A. H. Murray (Ed.), *Antisense RNA and DNA*; Wiley-Liss: New York, **1992**. [6b]E. Uhlmann, A. Peyman, *Chem. Rev.* **1990**, *90*, 543−584. [6c] U. Englisch, D. H. Gauss, *Angew. Chem. Int. Ed. Engl.* **1991**, *30*, 613−629; *Angew. Chem.* **1991**, *103*, 629−645.
[7] [7a] R. H. Hall, *The Modified Nucleoside in Nucleic Acid*; Columbia University Press: New York, **1971**. [7b]P. A. Limbach, P. F. Crain, J. A. AcCloskey, *Nucleic Acids Res.* **1994**, *22*, 2183−2196.
[8] [8a] S.-i. Kawahara, T. Wada, S. Kawauchi, T. Uchimaru, M. Sekine, *J. Phys. Chem. A* **1999**, *103*, 8516−8523. [8b] S.-i. Kawahara, T. Uchimaru, M. Sekine, *J. Mol. Struct. [THEOCHEM]* **2000**, *530*, 109−117. [8c] S.-i. Kawahara, T. Uchimaru, K. Taira, M. Sekine, *J. Phys. Chem. A* **2001**, *105*, 3894−3898. [8d] S.-i. Kawahara, A. Kobori, M. Sekine, K. Taira, T. Uchimaru, *J. Phys. Chem. A* **2001**, *105*, 10596−10611. [8e] S.-i. Kawahara, M. Sekine, K. Taira, T. Uchimaru, *J. Phys. Chem. A* **2002**, *106*, 3207−3212. [8f] S.-i. Kawahara, T. Uchimaru, *J. Mol. Struct. [THEOCHEM]* **2002**, *588*, 29−35.
[9] For a comprehensive review see: [9a] J. Sponer, P. Hobza, *Chem. Rev.* **1999**, *99*, 3247−3276. [9b] S. Scheiner, C. W. Kern, *Chem. Phys. Lett.* **1978**, *57*, 331−333. [9c] S. Scheiner, C. W. Kern, *J. Am. Chem. Soc.* **1979**, *101*, 4081−4085. [9d] P. Hobza, C. Sandorfy, *J. Am. Chem. Soc.* **1987**, *109*, 1302−1307. [9e] K. I. Trollope, I. R. Gould, I. H. Hiller, *Chem. Phys. Lett.* **1993**, *209*, 113−116. [9f] A. Destexhe, J. Smets, L. Adamowicz, G. Maes, *J. Phys. Chem.* **1994**, *98*, 1506−1514. [9g] I. R. Gould, P. A. Kollman, *J. Am. Chem. Soc.* **1994**, *116*, 2493−2499. [9h] P. Hobza, J. Sponer, M. Polásek, *J. Am. Chem. Soc.* **1995**, *117*, 792−798. [9i] J. Sponer, J. Leszczynski, P. Hobza, *J. Biomol. Strct. Dyn.* **1996**, *14*, 117−135. [9j] J. Sponer, P. Hobza, *Chem. Phys.* **1996**, *204*, 365−372. [9k] J. Sponer, P. Hobza, *Chem. Phys. Lett.* **1996**, *261*, 379−384. [9l] J. Sponer, J. Leszczynski, P. Hobza, *J. Phys. Chem.* **1996**, *100*, 1965−1974. [9m] J. Sponer, P. Hobza, *J. Phys. Chem. A* **2000**, *104*, 4592−4597. [9n] C. F. Guerra, F. M. Bickelhaupt, J. G. Snijders, E. J. Baerends, *J. Am. Chem. Soc.* **2000**, *122*, 4117−4128. [9o] L. Joubert, P. L. A. Popelier, *Phys. Chem. Chem. Phys.* **2002**, *4*, 4353−4359. [9p] E. Cubero, F. J. Luque, M. Orozco, *J. Am. Chem. Soc.* **2001**, *123*, 12018−12025. [9q] L. A. Popelier, L. Joubert, *J. Am. Chem. Soc.* **2002**, *124*, 8725−8729. [9r] N. A. Richardson, S. S. Wesolowski, H. F. Schaefer III, *J. Am. Chem. Soc.* **2002**, *124*, 10163−10170.
[10] J. Sponer, P. Hobza, *Chem. Phys. Lett.* **1997**, *267*, 263−270.
[11] [11a] G. A. Peterson, A. Bennett, T. G. Tensfeldt, M. A. Al-Laham, W. A. Shirley, J. Mantzaris, *J. Chem. Phys.* **1988**, *89*, 2193−2218. [11b] G. A. Peterson, M. A. Al-Laham, *J. Chem. Phys.* **1991**, *94*, 6081−6090.
[12] X/A//Y/B shows energy evaluation at the level of theory X with basis set A at the geometry optimized at the level of theory Y with basis set B.
[13] S.-i. Kawahara, T. Uchimaru, *Phys. Chem. Chem. Phys.* **2000**, *2*, 2869−2872.
[14] S. F. Boys, F. Bernardi, *Mol. Phys.* **1970**, *19*, 553−566.
[15] *SPARTAN:* D. J. Deppmeir, A. J. Driessen, T. S. Hehere, W. J. Hehere, J. K. Johnson, P. E. Klunzinger, L. Lou, J. Yu, J. Baker, J. E. Carpenter, R. W. Dixon, S. S. Fielder, H. C. Johnson, S. D. Kahn, J. M. Leonard, W. J. Piero, Wavefunction, Inc., Irvine CA, **1998**.
[16] *GAUSSIAN 94:* M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, J. A. Pople; Gaussian, Inc., Pittsburgh PA, **1995**.
[17] On inspection of the DNA/RNA-duplex standard geometry structure, close contacts between the substituent at the 8-position in purine/6-position in pyrimidine and the backbone are obvious. Because of steric hindrance, the nucleic acid base orientations of most of the 8-substituted A^X and G^X, and also the 6-substituted U^X and C^X species, are *syn* in nucleotides.[26] Thus, modified bases possessing substituents at these positions

may not form a favorable hydrogen bonding base pair in the standard DNA/RNA-duplex. However, if the structure of the backbone is drastically altered, such as in PNA,[27] the substituent at the 8-position will not be unfavorable for base pair formation through hydrogen bonds.

[18] Hammett parameter: X = NO_2; s*m* = 0.71, s*p* = 0.81, X = $COCH_3$; s*m* = 0.50, s*p* = 0.38, X = F; s*m* = 0.34, s*p* = 0.06, X = H; s*m* = s*p* = 0.00, X = NH_2; s*m* = −0.09, s*p* = −0.57. [18a] L. P. Hammett, *Physical Organic Chemistry 2nd Ed.*; McGraw-Hill: New York, **1970**. [18b] O. Exner, *Correlation Analysis in Chemistry*; (Eds.: N. B. Chapman, J. Shorter), Plenum Press: New York, **1978**.

[19] [19a] W. L. Jorgensen, J. Pranata, *J. Am. Chem. Soc.* **1990**, *112*, 2008−2010. [19b] J. Pranata, S. G. Wierschke, W. L. Jorgensen, *J. Am. Chem. Soc.* **1991**, *113*, 2810−2819. On the basis of secondary interaction, the most stable complex should be obtained when the orientation of all the hydrogen bonds is the same, and on the other hand, the stability of the hydrogen bond complex should be the least negative when the hydrogen bonds are formed in alternate orientation.

[20] Y. Kyogoku, R. C. Lord, A. Rich, *Proc. Natl. Acad. Sci. U. S. A.* **1967**, *57*, 250−257. They speculated that, in addition to the Watson−Crick-type base pair, a Hoogsteen-type base pair is partially involved.

[21] A^8F as the model of 1-methyl-8-bromoadenine [A^8Br]: the electrostatic properties of A^8Br should be similar to those of A^8F, so A^8F was considered the model of A^8Br in this section.

[22] The experimental values [*k*] should be compared with the difference in free energy, but the discussion in this work does not include the entropic contribution at all. Since the structures of the adenine derivatives studied in this work resemble each other, the entropic effect is likely to remain almost constant. However, due to the rotational freedom of the amino group, the magnitude of entropic contribution to the base pair formation of A^6NMe and A^6Nfo will possibly be slightly different from that of other derivatives but the difference should not be very important.

[23] R. Ishikawa, C. Kojima, A. Ono, M. Kainosho, *Magn. Reson. Chem.* **2001**, *39*, S159−S165.

[24] [24a] J. D. Bain, C. Switzer, A. R. Chamberlin, S. A. Benner, *Nature* **1992**, *356*, 537−539. [24b] C. Switzer, S. E. Moroney, S. A. Benner, *J. Am. Chem. Soc.* **1989**, *111*, 8322−8323. [24c] J. A. Piccirilli, T. Krauch, S. E. Moroney, S. A. Benner, *Nature* **1990**, *343*, 33−37. [24d] J. Horlacher, M. Hottiger, V. N. Podust, U. Hubscher, S. A. Benner, *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 6329−6333. [24e] R. X. F. Ren, N. C. Chaudhuri, P. L. Paris, S. Rumney, E. T. Kool, *J. Am. Chem. Soc.* **1996**, *118*, 7671−7678. [24f] J. C. Morales E. T. Kool, *J. Am. Chem. Soc.* **1999**, *121*, 2323−2324. [24g] E. T. Kool, J. C. Morales K. M. Guckian, *Angew. Chem. Int. Ed.* **2000**, *39*, 990−1009. [24h] A. K. Ogawa, Y. Wu, D. L. McMinn, J. Liu, P. G. Schultz, F. E. Romesberg, *J. Am. Chem. Soc.* **2000**, *122*, 3274−3287. [24i] T. Ohtsuki, M. Kimoto, M. Ishikawa, T. Mitsui, I. Hirao, S. Yokoyama, *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 4922−4925.

[25] For a comprehensive review of genetic code expansion with "natural base pairs" see: [25a] V. W. Cornishi, D. Mendel, P. G. Schultz, *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 621−633. [25b] M. Sisido, T. Hohsaka, *Bull. Chem. Soc., Jpn.* **1999**, *72*, 1409−1425.

[26] W. Saenger, *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, Berlin, Heidelberg, Tokyo, **1984**; Chapter 4.

[27] [27a] P. E. Nielsen, M. Egholm, R. H. Berg, O. Buchardt, *Science* **1991**, *254*, 1497−1500. [27b] B. Hyrup, P. E. Nielsen, *Bioorg. & Med. Chem.* **1996**, *4*, 5−23. [27c] E. Uhlmann, A. Peyman, G. Breipohl, D. W. Will, *Angew. Chem. Int. Ed.* **1998**, *37*, 2796−2823; *Angew. Chem.* **1998**, *110*, 2954−2983.

Received January 10, 2003

