Designing Molecular Switches Based on DNA-Base Mispairing

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · OCTOBER 2010

Impact Factor: 3.3 · DOI: 10.1021/jp106732u · Source: PubMed

2 AUTHORS:

Jissy A K
Karlsruhe Institute of Technology
14 PUBLICATIONS **157** CITATIONS
SEE PROFILE

Ayan Datta
Indian Association for the Cultivation of Sci…
96 PUBLICATIONS **1,489** CITATIONS
SEE PROFILE

Designing Molecular Switches Based on DNA-Base Mispairing

A. K. Jissy and Ayan Datta*

School of Chemistry, Indian Institute of Science Education and Research Thiruvananthapuram, CET Campus, Thiruvananthapuram-695016, Kerala, India

Received: July 20, 2010; Revised Manuscript Received: September 25, 2010

Stabilization of unstable mispairs on protonation in a DNA sequence can result in a change in the sequence conformation. Such sequences are being actively used for the synthesis of pH-driven molecular switches that have applications in biological pH sensing. We have studied various conformations of different mispairs of bases and their protonated forms using density functional theory (DFT) at B3LYP/6-31+G(d) and M05-2X/6-31+G(d,p) levels. Both gas-phase and aqueous-phase calculations are reported. Solvent phase calculations were done using the PCM and the COSMO solvation model. Our results show that the criterion for the protonation of a particular base in a mispair is not just its higher proton affinity. The planarity of the structure is significantly important, and a planar structure is energetically preferred over a bent mispair. Our calculations also show that the stabilization gained through protonation for the A−C, A−G, and the C−C mispairs is substantial (∼20.0 kcal/mol); therefore, these are good candidates for pH-driven molecular switches.

Introduction

Deoxyribonucleic acid (DNA) molecules store the genetic heritage of living organisms.[1] DNA molecules could act as building blocks for assembling nanoscale materials and devices by exploiting the highly specific interactions that occur between the base pairs. Hence, DNA is very attractive for self-assembling nanostructures and the construction of molecular devices.[2] The behavior of a given sequence can be reliably predicted by computer programs using the available thermodynamic data for DNA systems. Nowadays, virtually any DNA sequence can be designed using computer modeling and can be converted into a real molecule through DNA synthesizer. There are enzymes that can cut DNA in a specific manner (restriction endonucleases)[3] or join two pieces together.[4] DNA also has high physicochemical stability, much higher than that of proteins, and thus DNA-based nanodevices can be made, used, and stored under a broad range of environmental conditions.

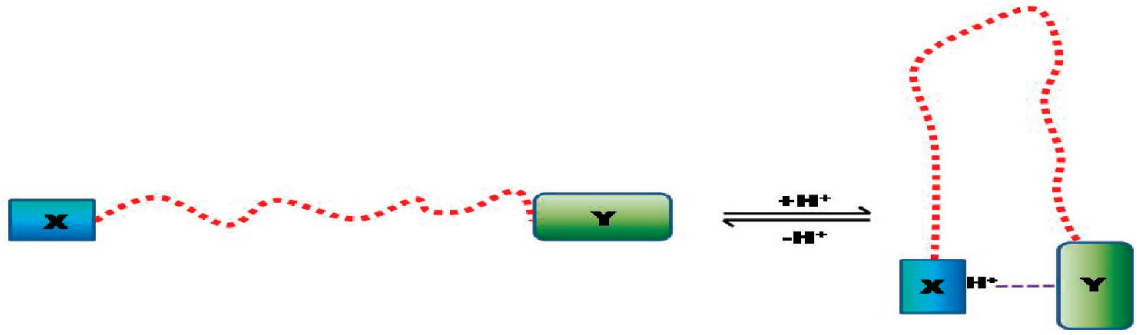
In DNA, almost exclusively, complementary nucleotides, that is, A···T and G···C, are incorporated during the process of nucleic acid biosynthesis. However, mispaired DNA bases can occur as a result of chemical/photochemical or physical damage to DNA, polymerase error during DNA replication, or recombination between nonhomologous parental DNA sequences.[5] In vivo errors due to the failure of the repair mechanism[6] after DNA replication occurs with a probability of 10^{-8} to 10^{-9} per base pair per replication.[7] The presence of such mismatches can lead to mutations if not corrected by proofreading or by the repair enzymes. Interestingly, the A···G mismatch is repaired less efficiently compared with the other mismatches.[8,9] However, some mismatches that are unstable are known to be stabilized significantly by the protonation of one of the bases under mildly acidic conditions.[10] The insertion of a single mispair can change the properties and the stability of DNA.[11] A DNA sequence that consists of a mispair can change its conformation as the pH is switched between neutral and slightly acidic. DNA nanomachines or assemblies capable of changing their conformation in response to an external stimulus like pH,[12,13] light, binding of a ligand,[14] electrochemical reaction,[15] or temperature and reversibly shift in between two stable states act as molecular switches and have a potential use in sensors, motors, and molecular computers.[16,17] Structural DNA nanotechnology involving design of molecular switches has a wide variety of applications.[18,19] Most of the DNA-based pH switches known until now use the protonation of cytosine to cause a transition between different structures.[19−21] Apart from the protonation of cytosine, recently, the stabilization of A−G mispair on protonation has also been used in the design of a pH-dependent molecular switch.[22] Such pH-triggered nanoswitches have fast response times, sustain efficiency over several cycles, produce nontoxic side products, and thus have some very useful in vivo as well as in vitro applications.[23] The possibility of using other non-Watson−Crick base pairs for the synthesis of molecular switches has not been considered. Until now, to the best of our knowledge, there has been no detailed theoretical study of the structures and the energetics of such base mispairing. The concept of a pH-driven nucleic acid switch is shown in Scheme 1.

Therefore, we have carried out high-level density functional theory (DFT) calculations for different conformers of isolated mispairs and their protonated structures in the gas phase as well as in the aqueous phase to have an atomistic detailed understanding about their potential uses in DNA nanotechnology.[24,25] We have studied the effect of protonation on the stability of free-standing mispairs without considering the sugar, phosphate groups, and other bases present in a DNA sequence. For a DNA sequence to function as a molecular switch, more than one mispair may be required to be inserted in the sequence[22] to provide the required stability on protonation. However, studying the structure and energetics at a single-base mispair level helps in understanding the molecular mechanism for such switching behavior. Also, such detailed quantum mechanical treatment will be the first step toward generating effective potentials for bulk molecular dynamics (MD) simulations for single strand → hairpin transformation. The conformational changes on protonation of the mispairs have been unravelled. Our calculations show that qualitatively acceptable predictions and quantitative

* To whom correspondence should be addressed. E-mail: ayan@iisertvm.ac.in.

10.1021/jp106732u
Published on Web 10/28/2010

SCHEME 1: Representation of a DNA Base Pairs Acting As a pH Switch[a]



[a] X,Y represents the nucleic acids: adenine, guanine, cytosine, thymine, or uracil.

estimation of thermodynamic binding energies for the base pairs held together by hydrogen bonds can be made. In the present study, the polarizable continuum model (PCM) is used for solvent studies,[26] which presents a very general numerical formulation allowing a mutual and self-consistent polarization between the solute charge density and the solvent.

Computational Methods

The structures for the different mispairs were obtained from the Protein Data Bank (PDB)[27] (PDB codes: 112D:[28] adenine····guanine; 420D:[29] protonated adenine····guanine; 113D:[30] guanine····thymine; 1D99:[31] cytosine····adenine; 1HWQ:[32] protonated adenine····cytosine; 1D80:[33] guanine····guanine; 438D:[34] guanine····uracil; 165D:[35] cytosine····uracil; 105D:[36] cytosine····cytosine; 225D:[37] protonated cytosine····cytosine; 283D:[38] adenine····adenine). The PDB files contain the entire DNA sequence from which we have extracted the particular mispair. The backbone containing the sugar, phosphate groups, and other base pairs was removed. Mispairs can exist in different configurations depending on the pH.[39] Considering this, in addition to the structures we obtained from the database, we have also generated the geometries of other possible mispairs in which at least two hydrogen bonds are formed in between the bases and protonation on nitrogen is possible. In all structures, normal −N−H···N or −N−H···O hydrogen bonds are formed, except for the mispairs that contain bifurcated H-bonds such as conformers of CC3 and H+CC1 as well as H+AC3 and GT3.

Molecular geometries and harmonic vibrational frequencies of the bases, the hydrogen bonded mispairs, and their protonated forms were obtained at the DFT[40] level. The nonlocal hybrid Becke three-parameter Lee−Yang−Parr (B3LYP),[41] as implemented in the Gaussian 03 program, was used for the energy minimization and frequency calculations with 6-31+g(d),[42] a split valence double-ζ standard basis set, augmented with a d-type polarization function on the non-hydrogen atoms. The convergence criteria for the geometry optimizations were set to less than 4.5×10^{-4} and 3.0×10^{-4} hartree/bohr for the maximum force and root-mean-square force per atom, respectively. The geometrical parameters and the total energies are given in the Supporting Information. We have considered only the protonation on nitrogen for the protonated structures because in the bases they are the positions with highest proton affinity.[43,44] The carbon atoms have very low basicities. There is no need to consider them for protonation because we desire to consider protonation of the mispair that stabilizes the base pair effectively for the sequence containing the particular mispair to be used as a molecular switch. Structural optimization was also performed at the M05-2X/6-31+G(d,p)[45,46] level. The M05-2X functional is known to be effective in incorporating long-range dispersion forces, so calculations at this level provide an estimate of such forces as well.[47] Dispersion forces are generally weaker than the forces associated with hydrogen bonding, but dispersion interactions play a major role in the stabilization of large biomolecules, such as DNA, in which these interactions are abundant. Anisimov et al.[48] have shown that dispersion makes a considerable contribution to the attractive forces in between the bases, thus enhancing their stability.

The interaction energies, ΔE, between the nucleic bases were evaluated as follows

$$\Delta E = E_{\text{basepair}} - E_{\text{A}} - E_{\text{B}}$$

where E_{basepair} is the total energy of the optimized base pair, whereas E_{A} and E_{B} are the total energies of the optimized individual bases, respectively. Finally, to obtain the corrected interaction energy, the basis set superposition error (BSSE) corrections were calculated using the counterpoise procedure.[49] The frequency calculations verified the absence of any vibrational instability in ground-state structures. To estimate the relative effect of water on the mispairs, we have used the PCM[50] on the B3LYP/6-31+G(d) level, as implemented in Gaussian 03,[51] and also used RI-DFT(COSMO) on the BP/def-TZVP level implemented in TURBOMOLE 6.0.[52] The resolution of identity (RI),[53] four center two integral approximation is used to reduce the computational cost without a loss in accuracy. In the PCM model of solvation, the solute is embedded in a cavity free of any solvent molecule, and the solvent (water) is considered to be a continuum outside the cavity. The cavity is determined from the molecular surface, as constructed from the scaled van der Waals radii of the constituent atoms. We have used cavities constructed by applying the united atom topological model, as implemented in the Gaussian code. The dielectric constant used for water is 78.39. The conductor-like screening model (COSMO)[54] method assumes conductor-like screening and empirically corrects for the effects of finite dielectric constant. It allows a flexible description of the solute charge distribution because it includes single center dipoles that are expected to be particularly important for centers with nonbonded electrons. COSMO also permits more rapid and efficient energy calculations and geometry optimizations for solutes described by general cavities. RI-DFTD calculations (geometry optimization) were also done at the BLYP/TZVPP level using TURBOMOLE to account for dispersion effects within the empirical correction formalism.[55] For the RI-DFTD and the COSMO calculations done in TURBOMOLE, the structures optimized by B3LYP/6-31+G(d) were used as initial guesses.

TABLE 1: Protonation Energies for Different Nucleic Bases

protonated base	protonation energy (kcal/mol)
H^+(N1) adenine	−18.9
H^+(N3) adenine	−17.4
H^+(N7) adenine	−10.5
H^+(N1) guanine	−4.5
H^+(N7)vguanine	−22.6
H^+(N3) cytosine	−22.0

Results and Discussion

The proton affinities (NIST database)[56] of the different bases are known to be in the following order

guanine > cytosine > adenine > thymine > uracil

We optimized the structures for the bases and their protonated structures at the B3LYP/6-31+G(d) level. The protonation energy for the different bases was evaluated considering the following reaction

$$\text{base} + NH_4^+ \rightarrow \text{base-}H^+ + NH_3$$

The results are shown in Table 1. The N7 position of guanine (Figure 1) has the highest basicity, which clearly indicates that during the protonation of an A···G mispair it is expected that guanine will get protonated. However, the observation is not as expected.[10b] Instead of guanine, adenine is the base that gets protonated upon lowering the pH of the system. Therefore, the structure of the A···G mispair was studied considering the protonation of adenine as well as guanine. The binding energy for the A···G mispair on protonation of adenine and guanine in the gas phase is −37.2 and −12.1 kcal/mol, respectively, thus creating a difference of ∼25.0 kcal/mol in the stability of the two structures. The structures of the mispair with protonated adenine and guanine are shown in Figure 2. The structures reveal that during the protonation of adenine, the planarity of the structure is maintained, whereas on protonation of guanine, the structure of the isolated mispair gets bent. The dihedral angle between atoms 1, 2, 3, and 4 (Figure 2) is 0.37° when adenine is protonated, whereas it is 76.4° on protonation of guanine. Such structural deformation reduces the stability of the guanine

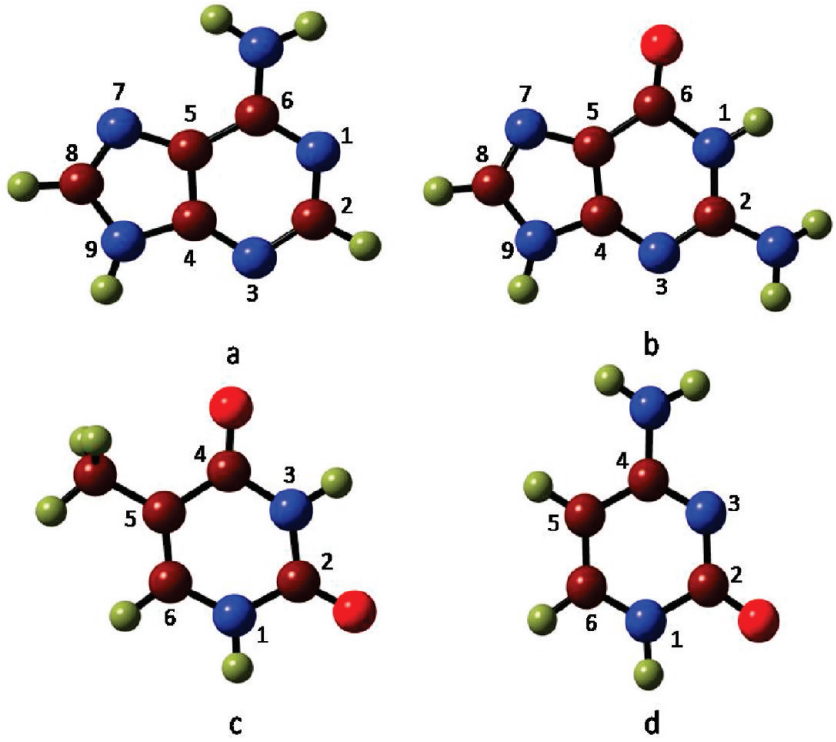


Figure 1. Structures of purine bases (a) adenine and (b) guanine and pyrimidine bases (c) thymine (uracil is without the methyl group) and (d) cytosine with conventional numbering scheme.

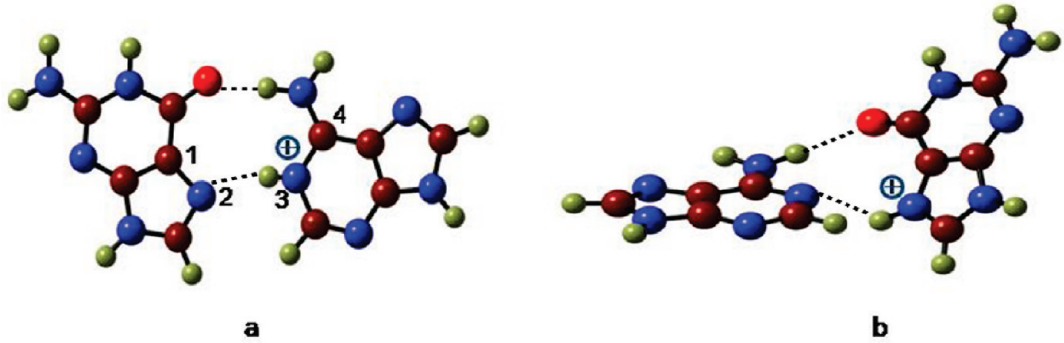


Figure 2. Structures of (a) planar H+A···G and (b) bent H+G···A.

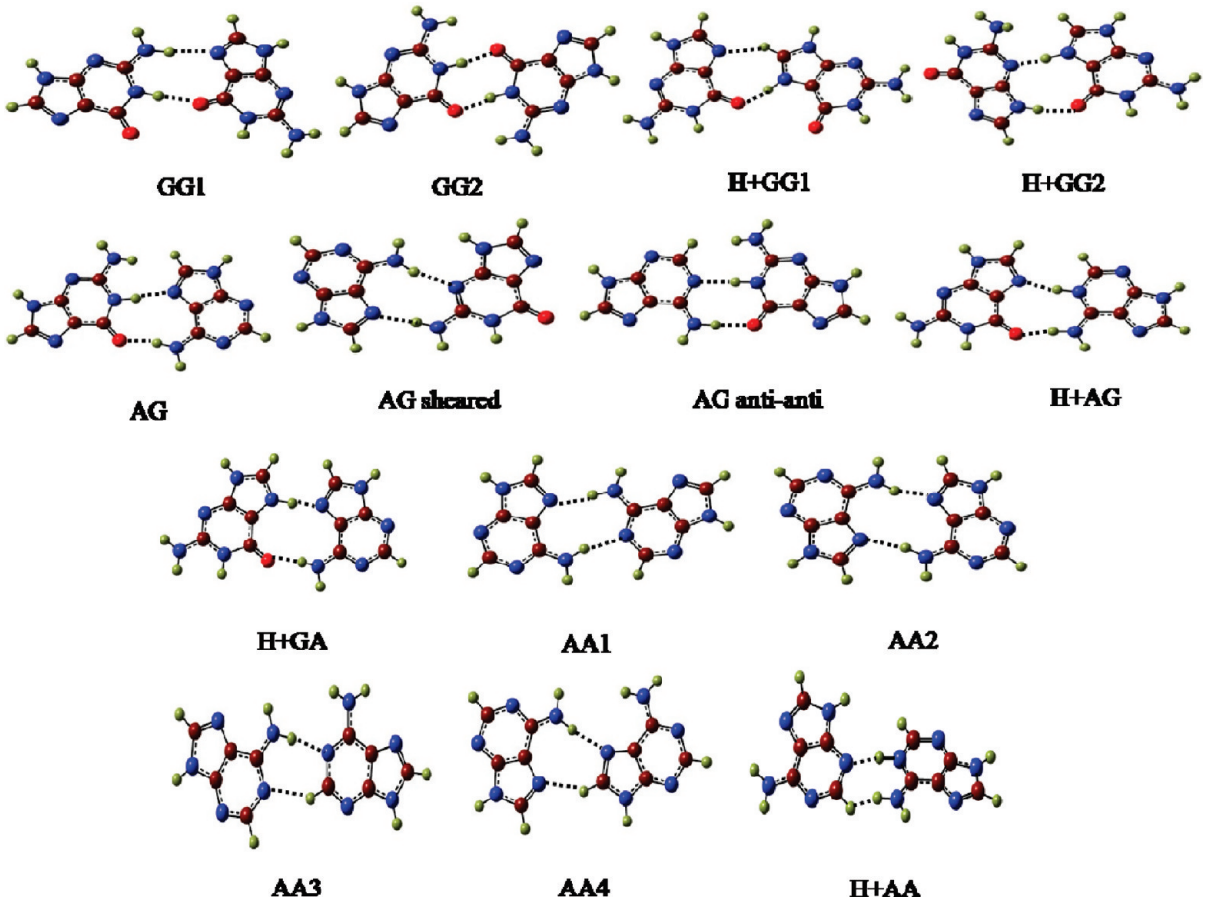


Figure 3. Optimized structures of purine−purine mispairs and their protonated forms.

protonated mispair. Therefore, the binding energies include the deformation energies required to transform the structures of the isolated bases into the structures they have in the isolated base pairs. In addition, such bending also leads to elongation of the H bonds. The N−H····O and N−H····N hydrogen bond lengths are 2.1; 2.3 and 1.7; and 2.0 Å for H+GA and H+AG, respectively. In DNA, the base pairs are stacked over each other. The planarity of the base pairs is an important criterion for such stacking. The base pairs would be associated with the sugar, phosphate group, and the backbone in DNA, and thus the system would be much more constrained than an isolated mispair. In such a case, the bending of the mispair on protonation and its lower stability will profoundly affect the formation of the hairpin loop. Therefore, the planarity-driven protonation of adenine would be favored in the mis-pair for intact DNA as well. This is applicable for this particular conformation of guanine (anti-syn) in which the N7-hydrogen on guanine tries to form a H-bond with N1 of adenine, which for adenine is the most basic position (Figure 2b). We can consider the protonation of A····G mispair for another conformation in which guanine forms a hydrogen bond with N7 of adenine (Figure 3, H+GA), which is its least basic nitrogen and therefore has a lesser tendency to take up the proton.

There are several other possible mispairs that could be made by different combinations of purines and pyrimidines. We have studied different conformations of these mispairs: purine−purine mispairs (Figure 3), pyrimidine−pyrimidine mispairs (Figure 4), and purine−pyrimidine mispairs (Figure 5) and their protonated forms to consider their potential use in the design of molecular switches. We have not studied the T···T, T···U, and U···U mispairs because in these three mispairs, none of the nitrogen atoms are free to take up a proton. We are considering only the protonation of nitrogen because it is the most basic and preferable site of attack. The higher the stability gained through protonation, the better the chance of the insertion of such mispair resulting in switching behavior on change of pH. If a mispair by itself is not stable but is considerably stabilized on protonation under slightly acidic conditions, then depending on the extent of stabilization that occurs, a DNA sequence containing the particular mispair would change its conformation on varying the pH from a single strand structure at neutral pH to a hairpin structure under slightly acidic conditions.[22]

The results of DFT/B3LYP calculation in the gas phase and in water and M05-2X calculations in the gas phase using Gaussian 03 and the results of RI-DFTD and COSMO calculations done using TURBOMOLE for the non-Watson−Crick pairs are reported in Table 2. The enthalpies for formation (ΔH) are also reported. The results for the three mispairs that show maximum stabilization on protonation are shown in bold.

Among the different structures studied, for the application as a molecular switch within intact DNA, AC1, AC2, H+CA2, H+AC2, CC1, CC2, H+CC1, H+GG2, GT2, and GU2 cannot be considered. In these, one of the hydrogen bonds formed is with the hydrogen, which in a DNA would be blocked by a sugar moiety. The M05-2X and RI-DFTD calculations predict a higher stability for all structures compared with the B3LYP results. This shows the significance of considering the additional attractive dispersive forces in the calculations.

CC3, AG sheared, GT3, and all AA mispairs are observed to be comparatively less stable among the mispairs studied, with a binding energy of <10.0 kcal/mol. The instability of CC3 can

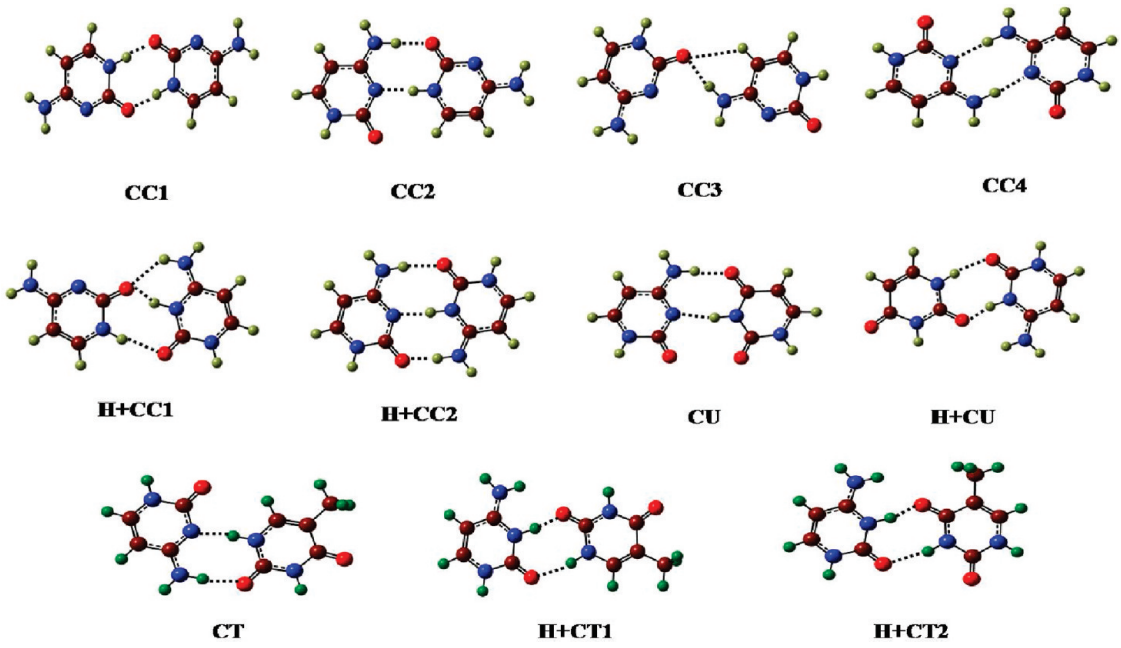


Figure 4. Optimized structures of pyrimidine−pyrimidine mispairs and their protonated forms.

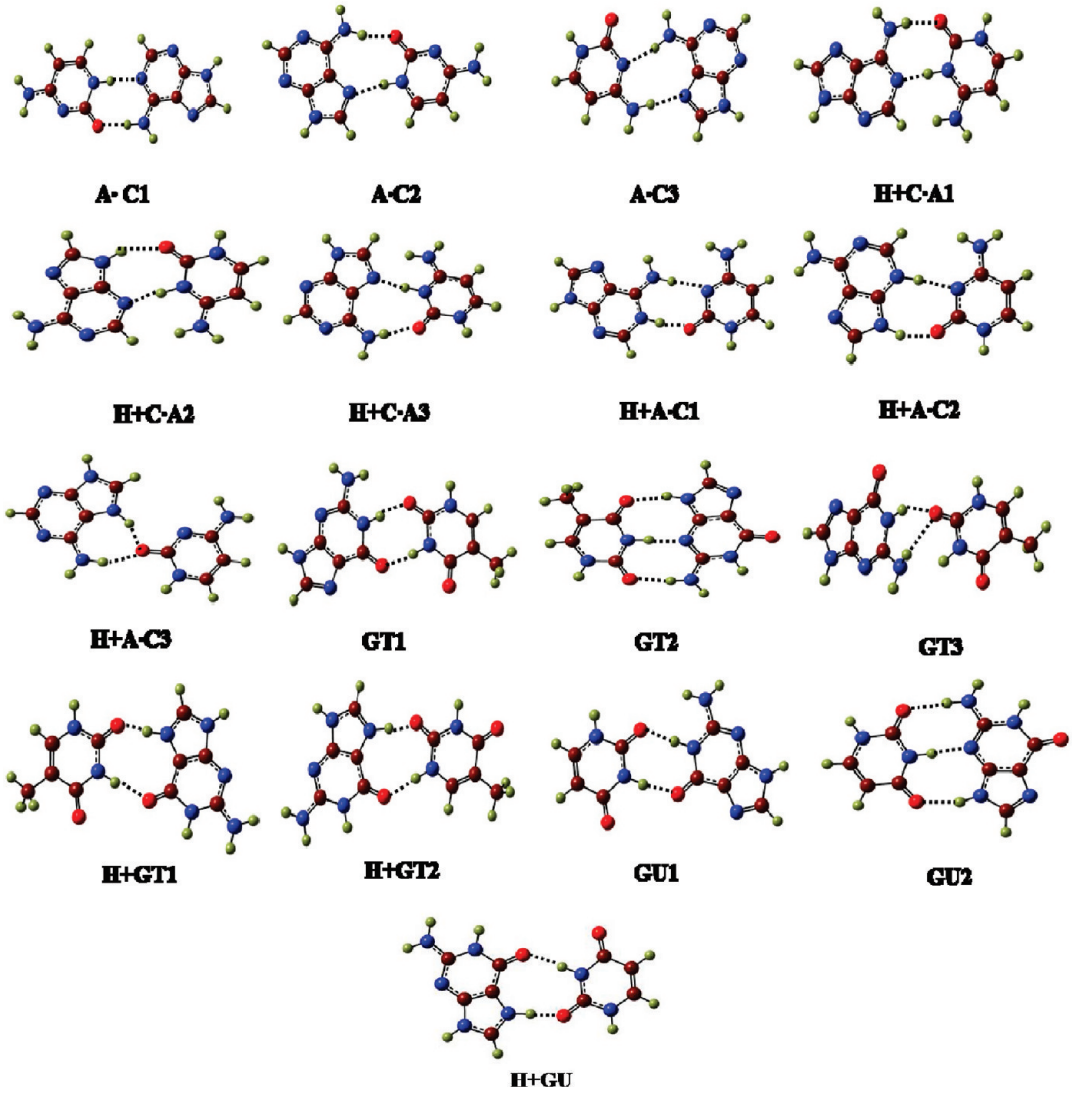


Figure 5. Optimized structures of purine−pyrimidine mispairs and their protonated forms.

TABLE 2: Binding Energies and Enthalpy of Formation (ΔH_f) for the Different Conformations of Base Pairs and Their Protonated Forms at Different Levels of Theory[a]

base pair	B3LYP/631+G(d) (gas phase)[b]	ΔH_f B3LYP/631+G(d) (gas phase)	M05-2X/6-31+G(d,p)[b]	B3LYP/631+G(d) (water)	DFT-D BLYP/TZVPP (TURBOMOLE)	COSMO (water) BP/def-TZVP (TURBOMOLE)
AC1	−14.1(−15.0)	−13.3	−15.6(−16.4)	−2.6	−18.5	−5.6
AC2	−13.0(−14.0)	−12.2	−14.2(−14.9)	−2.9	−16.9	−5.5
AC3	−12.4(−13.4)	−11.7	−14.1(−14.9)	−2.8	−17.3	−4.5
H^+CA1	−21.9(−22.9)	−22.1	−25.1(−26.1)	−0.3	−28.0	−4.5
H^+CA2	−22.4(−23.6)	−22.7	−24.5(−25.4)	1.2	−27.4	−4.9
H^+CA3	−18.3(−19.5)	−18.3	−20.8(−21.8)	−0.3	−25.0	−4.7
H^+AC1	−36.6(−37.7)	−35.9	−39.0(−39.9)	−5.1	−39.9	−9.1
H^+AC2	−34.8(−35.8)	−34.1	−38.2(−39.0)	−2.0	−38.6	−7.6
H^+AC3	**−37.7(−38.8)**	**−37.1**	**−40.4(−41.3)**	**−4.9**	**−40.0**	**−7.9**
CC1	−20.4(−21.4)	−19.7	−22.3(−23.2)	−2.9	−23.2	−6.6
CC2	−19.6(−20.7)	−19.0	−22.0(−22.9)	−2.8	−23.7	−5.9
CC3	−9.3(−9.9)	−8.3	−11.2(−11.7)	−1.8	−11.0	−2.0
CC4	−17.5(−18.6)	−16.9	−19.8(−20.6)	−2.7	−22.6	−4.2
H^+CC1	−34.9(−36.0)	−34.9	−39.6(−40.6)	−1.8	−39.0	−6.9
H^+CC2	**−42.5(−44.0)**	**−42.2**	**−45.5(−46.7)**	**−6.5**	**−47.8**	**−11.5**
CT	−18.5(−19.6)	−17.9	−21.3(−21.9)	−3.3	−22.7	−5.7
H+CT1	−23.7(−24.8)	−23.6	−26.9(−27.8)	−0.9	−27.9	−4.8
H+CT2	−24.9(−25.9)	−24.6	−27.6(−28.5)	−0.8	−29.2	−5.0
GG1	−12.0(−12.9)	−11.3	−26.8(−27.6)	−2.1	−12.5	−5.0
GG2	−23.4(−24.4)	−22.9	−26.7(−27.6)	−0.8	−28.0	−5.9
H^+GG1	−32.0(−33.0)	−31.8	−34.3(−35.0)	−3.0	−35.4	−5.6
H^+GG2	−15.4(−16.4)	−15.8	−18.1(−19.2)	−0.3	−26.9	−4.8
AG	−12.9(−14.0)	−12.2	−15.1(−15.9)	−1.3	−18.1	−4.7
AG sheared	−8.3(−9.2)	−7.8	−9.7(−10.7)	−2.1	−13.4	−3.7
AG anti-anti	−13.7(−14.6)	−13.0	−16.1(−16.8)	−0.9	−19.4	−4.5
H^+AG	**−37.2(−38.2)**	**−36.6**	**−38.9(−39.7)**	**−5.3**	**−41.5**	**−9.1**
H^+GA	−19.0(−20.1)	−19.4	−21.1(−22.1)	−2.9	−28.0	−6.3
H^+CU	−22.5(−25.4)	−22.3	−25.5(−26.4)	0.3	−26.5	−4.6
CU	−10.4(−11.4)	−10.0	−12.5(−13.3)	−0.2	−15.1	−3.2
GT1	−13.4(−14.4)	−13.0	−15.7(−16.6)	−1.2	−17.1	−4.7
GT2	−16.2(−17.5)	−16.0	−19.0(−20.0)	−2.2	−21.8	−6.3
GT3	−8.3(−9.2)	−7.5	−11.5(−12.3)	1.5	−12.5	−1.2
H^+GT1	−22.5(−23.6)	−22.6	−24.9(−25.9)	−2.5	−25.9	−5.8
H^+GT2	−23.0(−24.1)	−23.2	−25.7(−26.7)	−3.0	−26.6	−6.3
GU1	−13.3(−14.1)	−12.7	−15.3(−16.1)	−0.3	−16.8	−4.6
GU2	−16.1(−17.3)	−15.8	−18.8(−19.7)	−1.2	−21.6	−6.3
H^+GU	−21.4(−22.5)	−21.5	−23.7(−24.6)	−1.5	−24.7	−5.6
AA1	−9.4(−10.3)	−8.6	−10.5(−11.2)	−3.0	−14.1	−4.6
AA2	−8.1(−9.0)	−7.3	−9.4(−10.0)	−2.4	−12.8	−3.9
AA3	−4.6(−5.2)	−3.6	−6.3(−6.7)	−0.2	−8.2	−1.13
AA4	−5.6(−6.3)	−4.7	−6.9(−7.5)	−0.2	−9.2	−0.9
H^+AA	−20.0(−21.0)	−19.8	−22.1(−22.8)	−1.3	−25.4	−4.1
AT	−13.1(−14.1)	−12.4	−14.7(−15.5)	−3.0	−18.0	−5.2
GC	−24.7(−26.0)	−24.2	−27.7(−28.6)	−5.0	−29.7	−9.2

[a] All reported values are in kilocalories per mole. [b] BSSE-corrected (uncorrected) values are reported.

be attributed to the longer (∼2.9 Å) hydrogen bond by which it is stabilized in comparison with the H bonds (∼2.1 Å) in other mispairs. The dihedral angle between the planes of the rings for AG sheared, GT3, and AA2 are 53.8, 80.5, and 63.7°, respectively. Therefore, these mispairs do not possess a planar geometry, and hence bending is one of the reasons for the lesser stability of these mispairs. In addition, for the AG sheared and all four AA mispairs, both of the H-bonds are of N−H····N type, whereas in all other mispairs, there is one N−H····N- and one N−H····O-type hydrogen bond. The H-bond being formed with a lesser electronegative nitrogen further lowers the stability of these mispairs. Among the AA mispairs themselves, AA3 and AA4 have lower stability. The structures are planar but are stabilized by longer H-bonds. The N−H····N hydrogen bond lengths for AA2 and AA3 are 2.6 and 2.4 Å, respectively, whereas, it is <2 Å for other planar base pairs.

In the A····C mispair, irrespective of cytosine having a higher proton affinity as compared with adenine, the protonation of adenine leads to a much larger stabilization. This can be again explained on the basis of the planarity of the structures. In H+CA1 and H+CA3, the dihedral angles between the planes of the rings are −38.6 and −55.9°, respectively whereas both H+AC1 and H+AC3 are planar. On the protonation of cytosine, the planarity of the mispair is compromised, and the hydrogen bonds become longer. From the above discussion, it is clear that departure from a planar geometry in general lowers the stability of mispairs and reduces its ability to exhibit switching behavior on pH change.

Hobza et al.[57] determined accurate interaction energies for various nucleic acid pairs using RI-MP2 with aug-cc-pVQZ basis set. The interaction energy values they obtained for Watson−Crick pairs G−C and A−T are −27.7 and −15.1 kcal/mol, respectively. Anisimov et al.[58] also optimized isolated G−C and A−T canonical Watson−Crick DNA base pairs by the second-order Møller−Plesset (MP2) using different basis sets. The obtained interaction energies of the DNA bases vary from

TABLE 3: Stabilization Energy δ for Different Mispairs

mispair	δ (kcal/mol)
H^+AC	−23.6
H^+CA	−8.3
H^+CC	−22.1
H^+GG	−8.6
H^+AG	−23.5
H^+CU	−12.1
H^+GT	−6.8
H^+GU	−5.3
H^+AA	−11.2

−25.29 to −28.88 (G−C) and from −12.34 to −14.94 (A−T) kcal/mol. To check the reliability of our results, we also carried out calculations for the Watson−Crick G−C and A−T base pairs. The interaction energies obtained are −27.7 and −14.7 kcal/mol with DFT calculations at the M05-2X/6-31+G(d,p) level for the G−C and A−T, respectively, which compares quite well with the previously reported values.

In the case of solvated systems for all mispairs, it is observed that for PCM as well as the COSMO calculations, the stabilization of the base pairs is considerably reduced in water compared with the gas-phase structures (by at least 50%). This is expected because a polar solvent will solvate two separated polar bases better than the comparatively less polar dimer. H+AC3 is observed to be more stable in the gas phase in comparison with all other protonated AC mispairs, but in presence of water, H+AC1 is more stable than H+AC3. However, the role of solvation is less pronounced in reality because the water molecules are repelled by the hydrophobic sugar residues across the stacks of the base pairs. Therefore, gas-phase calculations can qualitatively predict the trends of binding of the base pairs in crystals or artificial assemblies.

The relative energies of the mispairs on protonation are significant. Therefore, after the optimization, the most stable structure for each mispair and its protonated form was considered to compare them with the nonprotonated mispair, respectively. The δ value is defined as

$$\delta = E_{\text{mis.prot.}} - E_{\text{mis}}$$

where $E_{\text{mis.prot}}$ and E_{mis} are the binding energies of the most stable conformations of protonated and nonprotonated mispair, respectively. The values are tabulated in Table 3. The total stabilization energies of the base pairs range from −23.6 to −5.3 kcal/mol. The maximum stabilization on protonation occurs for the A−C and the A−G mispairs, followed by the C−C mispair. Very interestingly, two of these three mis-pairs are in the literature as efficient pH-driven molecular switches. The use of the protonation of a C−C mispair in an i-motif for the construction of molecular switches is already reported.[59,60] The i-motif made up of cytosine-rich strands containing the $C-C^+$ mispair is known to work as an efficient pH sensor in the range of 5.0−7.0. DNA sequences containing two A−G mispairs also have been recently shown to work as an efficient reversible pH switch.[22] Our study shows that the H^+AC mispair formation can also be exploited for the construction of pH-dependent molecular switches because the stabilization provided by protonation in an A−C mispair is comparable to that for the C−C and the A−G mispairs.

Conclusions

Different physiological and pathogenic processes[61−63] can be elucidated if we could know about the changes occurring in the pH inside the cell. These intracellular changes could be characterized depending on the development of new biological sensors. Measurements of pH in humans indicate a difference between tumor and normal tissues.

DNA nanotechnology that has been developed on the structure of an i-motif containing the $C\cdots C^+$ mispair is elegant, but there is still a need to find new molecular switches driven on the change in pH, which can be more easily used because the i-motif requires many tracts of cytosine to be inserted into a structure, which limits its applications in certain assemblies. In our study, primarily, it is observed that irrespective of the fact that both guanine and cytosine have higher proton affinities compared with adenine, but still on protonation of the A−G and the A−C mispair, the stabilization is observed to be more if adenine is protonated. We suggest that the planarity of base pairs on protonation is an important parameter in the design of molecular switches. For applications in DNA molecular switch synthesis, we show that the stabilization obtained by protonation for the G−T, C−U, G−U, and the G−G mispairs is not sufficient enough for using them for the formation of a pH-driven molecular switch. The A−C mispair shows a large stabilization on protonation, which is comparable to the A−G mispairs already reported for pH-dependent switches. It has been experimentally shown that the structure of the A−C mispair varies as a function of the solvent pH.[39a] The mispair switches in between two conformations: the unstable reverse wobble at high pH and a much more stable wobble conformation at low pH. Therefore, the experimental evidence further validates the feasibility of the construction of a molecular switch based on the A−C mispair. The importance of the sequence context as well as the effect of the mispairs on different DNA secondary structures need to be examined to design an ideal pH switch based on the A−C mispair. Also, the role of cooperative interactions in stacks of A−C mispairs in hairpin formations needs further study.

Acknowledgment. We thank DST − Fast Track scheme for partial funding and University Grants Commission (UGC) for funding. We thank Prof. M. Hariharan for fruitful discussions.

Supporting Information Available: Cartesian coordinates for ground state structures, their energies in hartrees, and the harmonic frequencies. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Jacob, F. *The Logic of Life: A History of Heredity*; Princeton University Press: Princeton, NJ, 1993.

(2) (a) Modi, S.; Bhatia, D.; Simmel, F. C.; Krishnan, Y. *J. Phys. Chem. Lett.* **2010**, *1*, 1994. (b) Mohan, P. J.; Datta, A.; Sairam, S. M.; Pati, S. K. *J. Phys. Chem. B* **2006**, *110*, 18661. (c) Datta, A.; Pati, S. K. *Synth. Met.* **2005**, *155*, 398. (d) Sairam, S. M.; Pati, S. K. *J. Phys. Chem. B* **2007**, *111*, 11614.

(3) Kelly, T. J.; Smith, H. O. *J. Mol. Biol.* **1970**, *51*, 393.

(4) Mertz, J. E.; Davis, R. W. *Proc. Natl. Acad. Sci. U.S.A.* **1972**, *69*, 3448.

(5) Lindahl, T. *Nature* **1993**, *362*, 709.

(6) Iyer, R. R.; Pluciennik, A.; Burdett, V.; Modrich, P. L. *Chem. Rev.* **2006**, *106*, 302.

(7) Echols, H.; Goodman, M. F. *Annu. Rev. Biochem.* **1991**, *60*, 477.

(8) Marra, G.; Schär, P. *Biochem. J.* **1999**, *338*, 1.

(9) Leonard, G. A.; Booth, E. D.; Brown, T. *Nucleic Acids Res.* **2007**, *30*, 4740.

(10) (a) Kampf, G.; Kapinos, L. E.; Griesser, R.; Lippert, B.; Sigel, H. *J. Chem. Soc., Perkin Trans.* **2002**, *2*, 1320. (b) Carbonnaux, C.; van der Marel, G. A.; van Boom, J. H.; Guschlbauer, W.; Fazakerley, G. V. *Biochemistry* **1991**, *30*, 5449. (c) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.

(11) (a) Apalkov, V.; Berashevich, J.; Chakraborty, T. *J. Chem. Phys.* **2010**, *132*, 085102. (b) Hwang, G. T.; Hari, Y.; Romesberg, F. E. *Nucleic*

Acids Res. **2009**, *37*, 4757. (c) Natsume, T.; Ishikawa, Y.; Dedachi, K.; Tsukamoto, T.; Kurita, N. *Chem. Phys. Lett.* **2007**, *446*, 151.
(12) Wang, Z.; Lee, J. H.; Lu, Yi. *Chem. Commun.* **2008**, 6005.
(13) Chen, C.; Song, G.; Ren, J.; Qu, X. *Chem. Commun.* **2008**, 6149.
(14) Gradia, S.; Acharya, S.; Fichel, R. *J. Biol. Chem.* **2000**, *275*, 3922.
(15) Zazza, C.; Mancini, G.; Amadei, A.; Sanna, N.; Aschi, M. *Phys. Chem. Chem. Phys.* **2010**, *12*, 4552.
(16) (a) Bath, J.; Turberfield, A. J. *Nature Nanotechnol.* **2007**, *2*, 275. (b) Pitchiaya, S.; Krishnan, Y. *Chem. Soc. Rev.* **2006**, *35*, 1111.
(17) (a) Samori, B.; Zuccheri, G. *Angew. Chem., Int. Ed.* **2005**, *44*, 1166. (b) Seeman, N. C. *Biochemistry* **2003**, *42*, 7259.
(18) Liu, H.; Liu, D. *Chem. Commun.* **2009**, 2625.
(19) Chen, Y.; Lee, S.-H.; Mao, C. *Angew. Chem., Int. Ed.* **2004**, *116*, 5449.
(20) Ohmichi, T.; Kawamoto, Y.; Wu, P.; Miyoshi, D.; Karimata, H.; Sugimoto, N. *Biochemistry* **2005**, *44*, 7125.
(21) (a) Liu, D.; Balasubramarian, S. *Angew. Chem., Int. Ed.* **2003**, *42*, 5734. (b) Modi, S.; Swetha, D.; Goswami, M. G.; Gupta, G. D.; Major, S.; Krishnan, Y. *Nat. Nanotechnol.* **2009**, *4*, 325. (c) Wang, W.; Yang, Y.; Cheng, E.; Zhao, M.; Meng, H.; Liu, D.; Zhou, D. *Chem. Commun.* **2009**, 824. (d) Sharma, J.; Chhabra, R.; Yan, H.; Liu, Y. *Chem. Commun.* **2007**, 477.
(22) Lee, A.; De Rosa, M. C. *Chem. Commun.* **2010**, *46*, 418.
(23) Tannock, I. F.; Rotin, D. *Cancer Res.* **1989**, *49*, 4373.
(24) Sponer, J.; Leszczynski, J.; Hobza, P. *Biopolymers* **2001**, *61*, 3.
(25) Sponer, J.; Riley, K. E.; Hobza, P. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2595.
(26) Lipparini, F.; Scalmani, G.; Mennucci, B. *Phys. Chem. Chem. Phys.* **2009**, *11*, 11617.
(27) RCSB Protein Data Bank. http://www.pdb.org/pdb/home/home.do (accessed July 15, 2010).
(28) Brown, T.; Hunter, W. N.; Kneale, G.; Kennard, O. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 2402.
(29) Pan, B.; Mitra, S. N.; Sundaralingam, M. *Biochemistry* **1999**, *38*, 2826.
(30) Hunter, W. N.; Brown, T.; Kneale, G.; Anand, N. N.; Rabinovich, D.; Kennard, O. *J. Biol. Chem.* **1987**, *262*, 9962.
(31) Hunter, W. N.; Brown, T.; Kennard, O. *Nucleic Acids Res.* **1987**, *15*, 6589.
(32) Flinders, J.; Dieckmann, T. *J. Mol. Biol.* **2001**, *308*, 665.
(33) Skelly, J. V.; Edwards, K. J.; Jenkins, T. C.; Neidle, S. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 804.
(34) Pan, B.; Mitra, S. N.; Sundaralingam, M. *Biochemistry* **1999**, *38*, 2826.
(35) Cruse, W. B.; Saludjian, P.; Biala, E.; Strazewski, P.; Prange, T.; Kennard, O. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4160.
(36) Leroy, J. L.; Gueron, M. *Structure* **1995**, *3*, 101.
(37) Gehring, K.; Leroy, J. L.; Gueron, M. *Nature* **1993**, *363*, 561.
(38) Baeyens, K. J.; De Bondt, H. L.; Pardi, A.; Holbrook, S. R. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12851.
(39) (a) Boulard, Y.; Cognet, J. A. H.; Gabarro-Arpa, J.; Le Bret, M.; Sowers, L. C.; Fazakerley, G. V. *Nucleic Acids Res.* **1992**, *20*, 1933. (b) Sowers, L. C.; Boulard, Y.; Fazakerley, G. V. *Biochemistry* **2000**, *39*, 7613.
(40) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787.
(41) (a) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 78.
(42) Tian, S. X. *J. Chem. Phys. A* **2005**, *109*, 5153.
(43) Chen, X.; Syrstad, E. A.; Nguyen, M. T.; Gerbaux, P.; Tureček, F. *J. Phys. Chem. A* **2004**, *108*, 9283.
(44) Xia, F.; Xie, H.; Cao, Z. *Int. J. Quantum Chem.* **2008**, *108*, 57.
(45) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 5656.
(46) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2*, 364.
(47) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 289.
(48) Danilov, V. I.; Anisimov, A. M. *J. Biomol. Struct. Dyn.* **2005**, *22*, 471.
(49) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
(50) Miertus, S.; Sorocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117.
(51) 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.; Pople, J. A. *Gaussian 03*; Gaussian, Inc.: Wallingford, CT, 2004.
(52) (a) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165. (b) von Arnim, M.; Ahlrichs, R. *J. Comput. Chem.* **1998**, *19*, 1746.
(53) (a) Kendall, R. A.; Früchtl, H. A. *Theor. Chem. Acc.* **1997**, *97*, 158. (b) Vahtras, O. O.; Almlöf, J.; Feyereisen, M. W. *Chem. Phys. Lett.* **1993**, *213*, 514.
(54) Klamt, A.; Schüürmann, G. *J. Chem. Soc., Perkin Trans.* **1993**, *2*, 799.
(55) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463.
(56) Proton Affinity Search. http://webbook.nist.gov/chemistry/pa-ser.html (accessed July 15, 2010).
(57) Šponer, J.; Jurečka, P.; Hobza, P. *J. Am. Chem. Soc.* **2004**, *126*, 10142.
(58) Kurita, N.; Danilov, V. I.; Anisimov, V. M. *Chem. Phys. Lett.* **2005**, *404*, 164.
(59) Liedl, T.; Simmel, F. *Nano Lett.* **2005**, *5*, 1894.
(60) Alberti, P.; Bourdoncle, A.; Saccà, B.; Lacroix, L.; Mergny, J. L. *Org. Biomol. Chem.* **2006**, *4*, 3383.
(61) Matsuyama, S.; Llopis, J.; Deveraux, Q. L.; Tsien, R. Y.; Reed, J. C. *Nat. Cell Biol.* **2000**, *2*, 318.
(62) Mukherjee, S.; Ghosh, R. N.; Maxfield, F. R. *Physiol. Rev.* **1997**, *77*, 759.
(63) Montesano, R.; Roth, J.; Robert, A.; Orci, L. *Nature* **1982**, *296*, 651.

JP106732U