*J. Phys. Chem. A* **2003,** *107,* 9447−9455

# Isoguanine: From Base Pair to Tetrad

**Jiande Gu[†] and Jerzy Leszczynski*[,‡]**

*Drug Design & Discovery Center, State Key Laboratory of Drug Research,*
*Shanghai Institute of Materia Medica, Shanghai Institutes for Biological Sciences,*
*Chinese Academy of Sciences, Shanghai, 200031 P. R. China, and*
*Computational Center for Molecular Structure and Interactions,*
*Department of Chemistry, Jackson State University, Jackson, Mississippi 39217*

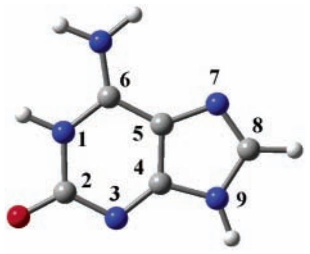
*Received: June 30, 2003*

Five different isoguanine (isoG) base pairs and five isoG tetrads have been located as the local minima on the potential energy surfaces at the B3LYP/6-311G(d,p) level of theory. An analysis of the isoG base pair enables us to evaluate the hydrogen bond energy in isoG−iosG interactions which is useful in the study of the energy properties of the isoG tetrad. All isoG tetrads in cyclic form are more stable than guanine tetrads. The difference in the stabilization energy between the planar and nonplanar conformers of the isoG tetrad is negligible. With the bond energy assignment, we are able to investigate the cooperative effect of hydrogen bonding in the cyclic tetrads. The cooperative effects of hydrogen bonding are crucial for the highly stabilized isoG tetrads. The balance between proton donation and the accepting of protons in the isoG tetrad governs the cooperative effects. Bifurcated H-bonding is important in the formation of isoG tetrads. They provide alternatives for isoG to develop various stable tetrad conformers which have the potential to create diversity in biological systems. The electrostatic potential (ESP) maps reveal that the nonplanarity of the isoG base pairs originates from the electrostatic repulsion between the monomers in the pair. ESP maps for isoG tetrads also indicate that placing a cation, especially alkali metal ions, in the central area of the isoG tetrads will inevitably lead to the bowl-like structure.

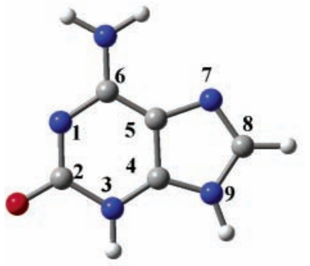
## Introduction

Isoguanine (2-oxo-6-amino-guanine, isoG, Figure 1), with an oxygen at the C2 position of purine, is an oxidized derivative of adenine. It has been found to occur naturally and may be considered an elementary nucleobase.[1,2] The physical characteristics of isoG resemble guanine, and isoG also forms tetraplexes. Similarly to the guanine tetrad, the higher-order self-pairing of isoG has been recognized.[1,3−7] NMR data suggest that the isoG mononucleoside aggregates to form octamers in organic solvents with the presence of several different cations.[4] The ion exchange HPLC experiments for d($T_4$isoG$_4T_4$) and d($T_4G_4T_4$) indicate that the isoG tetrads might be more stable than the G tetrads.[5] On the other hand, unlike the guanine tetrad which adopts either the Hoogsteen bases-pair motif [8] or the bifurcated hydrogen-bonding form,[9] several different conformers have been suggested for the isoG tetrads.[4−7,10,11] A bowl-shape structure in the presence of a potassium cation has been proposed by Davis et al.[4] based on the NMR data. The Hoogsteen bases-pair motif was suggested for aggregations of isoguanosine.[10] The Hoogsteen bases-pair mixed with single H-bonding between O2 and H(N6) has been portrayed for poly(isoG) tetraplexes.[6] In addition, tetrads consisting of mixtures of 1H-isoG and 3H-isoG have been postulated in pyranosyl RNA.[11]

Recently, isoG-contained DNA pentaplexes have been designed and characterized.[1,3,12] DNA pentaplexes are expected to be able to form ion channels because of their morphological similarity to the pentameric alpha-helical peptides, which have been implicated as ion channels.[13]



1H-isoG



3H-isoG

**Figure 1.** 2-Oxo-6-amino-guanine (isoG). Color representations for atoms are red for O, blue for N, gray for C, and white for H.

Base pairing is one of the most important factors in the formation of tetraplexes. An understanding of base pairing is the first step toward uncovering the stability and structural characteristics of base tetrads. Previous theoretical studies of

---

* To whom correspondence should be addressed.
[†] State Key Laboratory of Drug Research.
[‡] Jackson State University.

10.1021/jp030798j CCC: $25.00

different base tetrads revealed that the stable tetrads play important roles in the formation of stable tetraplexes.[3,4,11,15−23] Also, to understand the diversity of DNA tetraplexes, a knowledge of the relative stability of different conformers of the tetrad and their bonding characteristics is essential. The details of base pairing patterns and their relation to the relative stability of different conformers could only be revealed by accurate computational investigations.

For DNA related tetraplexes, the possibility for the existence of planar or near-planar tetrad structures is of special importance because they can be further stabilized by successive tetrads through stacking interactions. However, according to the schematic representation suggested by Chaput et al.,[1] the sector angle of 67° for isoG defined by the intersection of vectors along the van der Waals donor−donor and acceptor−acceptor atom surfaces is not favorable for a tetrad with the Hoogsteen bases-pair motif in planar or near planar structures.

In our recent efforts to understand the role of H-bonding in the formation of base tetrads, the bifurcated H-bonding pattern has been found to be very important for stabilizing guanine tetrads without the presence of cations.[8,9] Quantum chemical studies of the stability of guanine tetrads reveal that the binding energy of the G tetrad with the bifurcated H-bonding pattern is not lower than that in the Hoogsteen bases-pair motif.[8,9,12,13] While forming bifurcated H-bonding conformers, the bases are expected to be able to adjust the bonding angle to form planar or near-planar tetrads. Such complexes are suitable to intercalate between the successive tetrads in tetraplexes.

The symmetric form of the isoG tetrad ($C_{4h}$ symmetry) and the isoG quintet have been studied by Meyer and Suhnel at the B3LYP/DZVP level of theory.[24] However, their research has been focused on the interaction between cations and the isoG tetrad. The isoG quintet has also been studied at the B3LYP/6-311G(d,p) level of theory in our laboratory.[25] In this study, the cooperative effects in H-bonding[26] for the cyclic quintet of isoG system are found to be substantial. Here we report the results of a DFT-based quantum chemical study of the stability, molecular structures, H-bonding pattern, and electrostatic potential properties of the isoG tetrad. To analyze the base-pairing contribution to the formation of isoG tetrads, the properties of the base-pairs appearing in the different conformers of isoG4 are also discussed. The theoretical results presented in this paper provide new information concerning important features of nucleobase tetrads. Even though the study described here does not take into account interactions with cations, the calculated electrostatic potentials of the tetrad shed light on the possible hosting positions for metal ions.

## Calculation Methods

The local minima of the tetrad complexes have been fully optimized by analytic gradient techniques. The method used was the density functional theory (DFT) with Becke's three parameter (B3)[27] exchange functional along with the Lee−Yang−Parr (LYP) nonlocal correlation functional (B3LYP).[28,29] The standard 6-311G(d,p) valence triple-$\zeta$ basis set augmented with d-type and p-type polarization functions[30] was used for all elements. It is well-known that the geometries and vibrational frequencies of the molecules calculated at the B3LYP/6-311G(d,p) level agree well with experiment.[31,32] In the vibrational frequency calculations, the force constants were determined analytically. The Gaussian-98 package of programs[33] was used in the calculations. The basis set superposition error (BSSE) was corrected according to Boys' approach.[34]

## Results and Discussions

isoG can adopt two forms: 1H-isoG and 3H-isoG. The energy difference between the fully optimized structures of these two forms amounts to less than 0.4 kcal/mol at the B3LYP/6-311G(d,p) level of theory, with 1H-isoG being more stable. Like normal guanine, 1H-isoG exhibits the nonplanar amine group in its optimized structure.[35] However, 3H-isoG, as a local minimum on the potential energy surface, adopts the $C_s$ symmetry with a planar amino group. Both 1H-isoG and 3H-isoG could be involved in the formation of the base-pairs and tetrads of isoG.

**Base Pairs of isoG.** Six different prototypes of the isoG−isoG base pair, which appear in the isoG tetrads, were optimized at the B3LYP/6-311G(d,p) level of theory. Their structures are depicted in Figure 2 along with the main geometrical parameters. Among these base pairs, **d1** and **d2** consist of 1H-isoG and 3H-isoG and the others comprise only the 1H-isoG form. Vibrational frequency analysis indicates that, among the local minima on the potential energy surface, only **d1** adopts the planar structure.

Three H-bonds hold 1H-isoG and 3H-isoG tightly in the base pair **d1**. A similar bonding pattern has been established for the normal GC base pair.[17] The short H(N6)···O2′ bond distance of 1.727 Å in **d1** suggests that 3H-isoG, acting as a double H-bonding acceptor, tends to donate its proton at the N6 position to form an H-bond with the O2 of 1H-isoG. This allows for the compensation of the extra charge loss caused by receiving two protons for the formation of the H-bonds of N1···H3′ and O2···H(N6)′. With the similar argument, the relatively weaker O2···H(N6)′ bond is expected as is indicated by its slightly longer H-bond length of 1.880 Å. Consequently, the O2 of 3H-isoG and the nonbonded H(N6) of 1H-isoG in **d1** are expected to be involved in the formation of H-bonds with other nucleobases.

Comprising both the 1H-isoG and 3H-isoG forms, **d2** adopts the nonplanar structure. The dihedral angle of the four atoms involved in H-bonding (H(N6)−N7−H3′−O2′) amounts to 15.5°. The planar form is the transition state corresponding to the flipping of the bases as revealed through vibrational mode analysis. Two isoGs involved as the H-bonding donor and acceptor are held together by the H(N6)···O2′ and N7···H3′ bonds. The short bond lengths of 1.729 Å for the former and 1.906 Å for the latter indicate strong H-bonding in **d2**.

$C_2$ symmetry is assigned to the local minimum **d3** which consists of only the 1H-isoG form. Two isoGs tilt at 25.5° forming an H(N6)−N7−H(N6)′−N7′ dihedral angle in the base pair. The H(N6)−N7 atomic distance amounts to 1.939 Å, about 0.03 Å longer than that in **d2**. The H-bonds that hold the base pair are slightly weaker than those in **d2**, although isoG also acts as an H-bond donor and acceptor at the same time. The planar form with $C_{2h}$ symmetry has been revealed by vibrational frequency calculations to be the transition state structure. It links the two identical energy **d3** local minima. The H-bonds in the planar form of **d3** are slightly reduced compared to those in the energy minimum structure (1.960 vs 1.939 Å for the H(N6)−N7 bond length).

With one isoG as a double H-bonding donor and the other as a double H-bonding acceptor, two 1H-isoGs are bonded loosely through H(N6)···N3′ and H1···O2′ in **d4**. The length of the H-bond H(N6)···N3′ has been predicted to be 2.137 Å, about 0.20 Å longer than that of H(N6)···N7′ in **d3** (0.23 Å longer than that in **d2**). The bond length of H1···O2′ in **d4** has been evaluated to be 1.961 Å, also 0.23 Å longer than that in **d1** and **d2**. The value of 15.9° for the O2′−N3′−H(N6)−H1 dihedral angle in the energy minimum structure **d4** reflects the

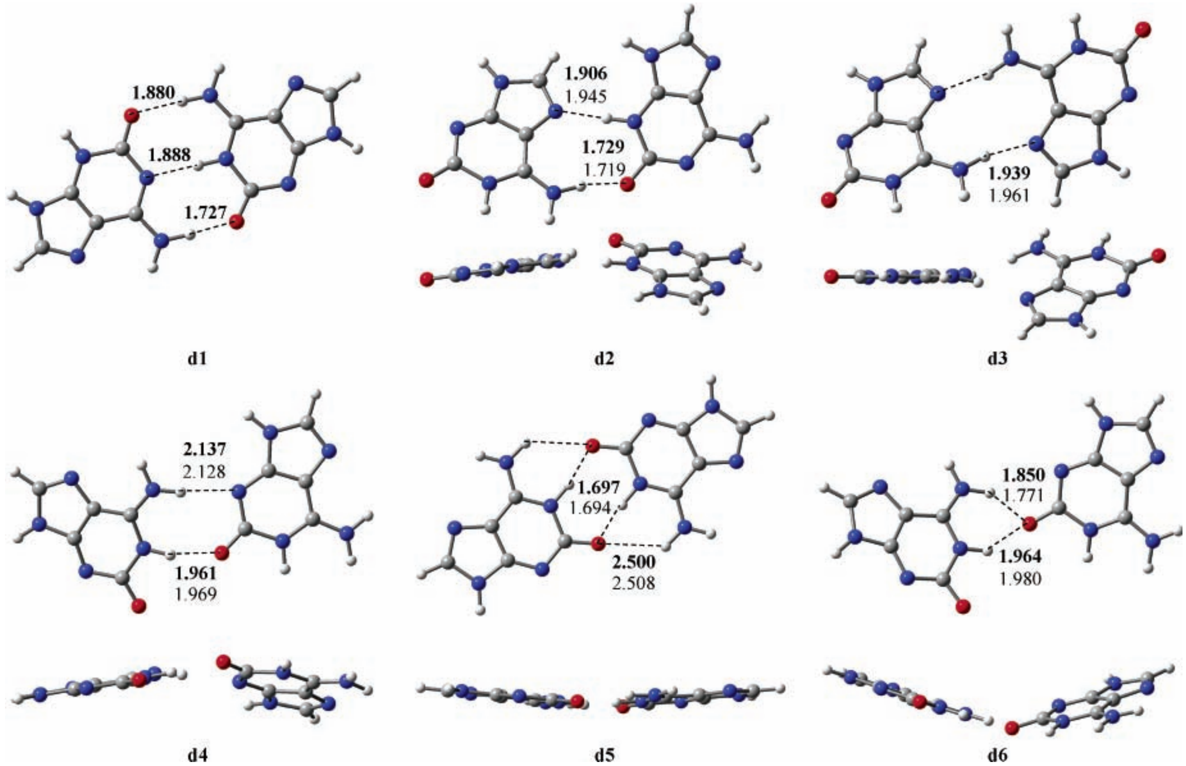


**Figure 2.** Optimized isoG base pairs with side views. The main geometric parameters are for the energy local minimal structures (bold) and for the planar structures (plain). **d6** is not the fully optimized structure.

**TABLE 1: Energy Properties of isoG Base Pairs Evaluated at the B3LYP/6-311G(d,p) Level[a]**

| | symmetry | composition | $E$/hartree | $\Delta E$/kcal mol$^{-1}$ | $E°$/hartree | $\Delta E°$/kcal mol$^{-1}$ | $\Delta E°$ (BSSE) |
|---|---|---|---|---|---|---|---|
| dimmer | | | | | | | |
| **d1** | $C_s$ | 1H-isoG: 3H-isoG | −1085.429516 | −30.50 | −1085.194733 | −28.54 | −25.71 |
| **d2** | $C_1$ | 1H-isoG: 3H-isoG | −1085.411275 | −19.05 | −1085.177665 | −17.83 | −15.00 |
| | $C_s$ | | −1085.410851 | −18.79 | −1085.177524 | −17.74 | −14.91 |
| **d3** | $C_2$ | 1H-isoG: 1H-isoG | −1085.403837 | −14.06 | −1085.169893 | −12.94 | −10.04 |
| | $C_{2h}$ | | −1085.402917 | −13.48 | −1085.169518 | −12.70 | −9.84 |
| **d4** | $C_1$ | 1H-isoG: 1H-isoG | −1085.405885 | −15.34 | −1085.172103 | −14.32 | −11.42 |
| | $C_s$ | | −1085.405800 | −15.29 | −1085.172490 | −14.56 | −11.66 |
| **d5** | $C_2$ | 1H-isoG: 1H-isoG | −1085.429700 | −30.29 | −1085.195432 | −28.96 | −26.06 |
| | $C_{2h}$ | | −1085.429693 | −30.28 | −1085.195588 | −29.06 | −26.16 |
| **d6** | $C_1$ | 1H-isoG: 1H-isoG | −1085.406760 | −15.90 | −1085.172600 | −14.68 | −11.78 |
| | $C_s$ | | −1085.406844 | −15.95 | −1085.173117 | −15.01 | −12.11 |
| monomer | | | | | | | |
| 1H-isoG | $C_1$ | | −542.6907167 | | −542.574618 | | 1.45* |
| 3H-isoG | $C_s$ | | −542.6901951 | | −542.574640 | | 1.38* |

[a] $E^0$ and $\Delta E^0$ are vibrational zero-point energy corrected, where $\Delta E = E(\text{dimmer}) - E(\text{monomer1}) - E(\text{monomer2})$, $\Delta E^0 = E^0(\text{dimmer}) - E^0(\text{monomer1}) - E^0(\text{monomer2})$, and $\Delta E^0(\text{BSSE}) = E^0(\text{dimmer}) - E^0(\text{monomer1}) - E^0(\text{monomer2}) + \text{BSSE}(\text{monomer1}) + \text{BSSE}(\text{monomer2})$. The asterisk is the BSSE of monomer. **d6** are the partially optimized calculation results.

influence of nonplanarity of the amine group in isoG. The planar structure in this base-pair shows the characteristics of a transition state between the two structure-like energy local minima as revealed by the vibrational frequency analysis.

The **d5** local minimum structure adopts an almost planar form. The H-bond length of O2···H1′ is as short as 1.697 Å, the shortest among all of the H-bonds predicted in this study. The atomic distance of 2.500 Å between O2 and H(N6)′ suggests the existence of a weak H-bond. In this sense, 1H-isoGs in **d5** are held by two pairs of bifurcated H-bonds. The slight reduction of the O2···H(N6)′ distance (by 0.003 Å) and the tiny elongation of O2···H1′ (by 0.008 Å) from the planar form to the nonplanar conformer implies that isoG is able to adjust in order to create a better bifurcated H-bonding pattern.

Base-pair pattern **d6** could not be obtained from unrestricted geometry optimization. The geometric parameters of H-bonds in both the planar and nonplanar forms were taken from the optimized structures of the isoG tetrads which will be discussed below. These geometric parameters of H-bonds were kept frozen during the partial optimization of the base pairs. The H bond presented in **d6** displays a typical bifurcated H-bond.

The energy properties of the isoG base pairs are summarized in Table 1. The energy difference between the local minima and their corresponding planar structures is less than 0.6 kcal/mol. This insignificant difference implies an easy flapping motion for the bases. The small values of the imaginary frequency of those planar transition state structures further justify this easy motion. With these energy data along with the geometries of the isoG base pairs, we are able to assign approximate bonding energies to different types of H-bonds that exist in isoG base pairs. Among the six base pairs, **d1** and **d5** are the most stable. The least stable base pair has been found to be **d3** in which two H···N bonds contribute about 10.0 kcal/mol (with BSSE and zero-point corrections) to stabilize this

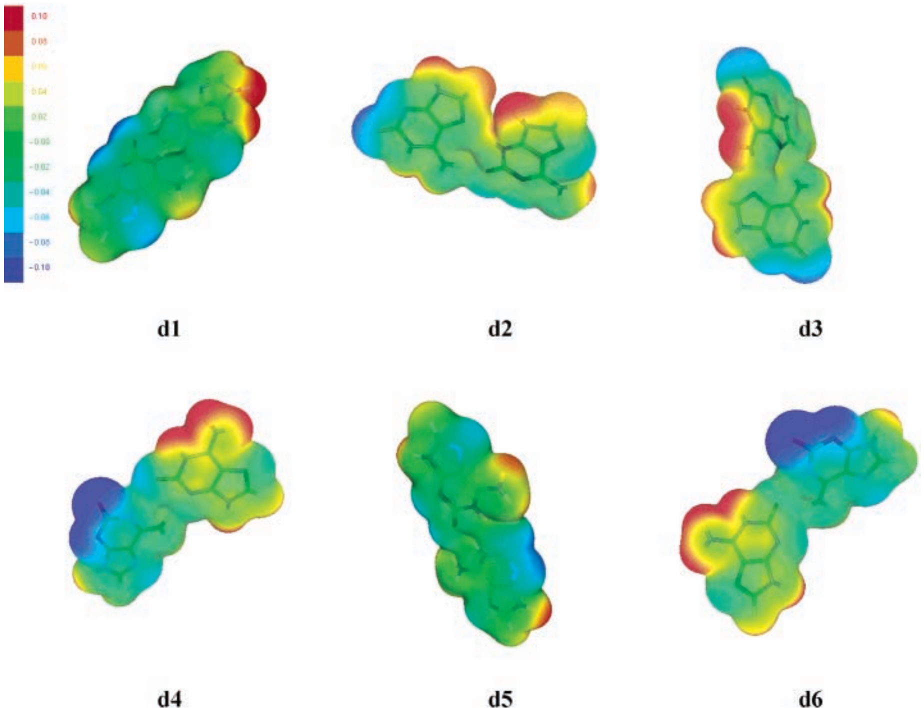


**Figure 3.** Electrostatic potential maps of the isoG base pairs. Electrostatic potential of isoG pairs are mapped onto the surface of the electron density of 0.002 unit.

dimer. In this way, the bonding energy for this H···N bond can be estimated to be approximately 5.0 kcal/mol. The stabilization energy in **d2** amounts to 15.0 kcal/mol at the level of the present study. Because **d2** contains one H···N bond with a similar bond length as that in **d3**, the 5.0 kcal/mol value of bonding energy could be a good approximation for the N7···H3′ bond in **d2**. Accordingly, the bonding energy of H···O is evaluated to be around 10.0 kcal/mol. The stabilization energy of 25.7 kcal/mol in **d1** is attributed to two H···O bonds and one H···N H-bond. A rough estimation based on the bonding energy in **d2** and **d3** yields a value of 25.0 kcal/mol for two H···O bonds and one H···N H-bond which is a reasonable approximation compared to 25.7 kcal/mol for the DFT level value. The similar H-bonding pattern and the stabilization energy of 25.4 kcal/mol can also be found in the GC base pair.[18] The structure in which one isoG acts as a double H-bonding donor and the other acts as a double H-bonding acceptor weakens the H-bonding in **d4**. The stabilization energy for **d4** amounts to 11.4 kcal/mol, about 3.6 kcal/mol lower than that in **d2**. The unbalanced charge distribution in **d4** reduces the stabilization energy. This corresponds to the elongation of the H-bond lengths as discussed above. Two bifurcated H-bonds have been identified in **d5**. The stabilization energy has been evaluated to be 26.06 kcal/mol for this conformer. The bonding energy of a bifurcated H-bond amounts to approximately 13.0 kcal/mol which is 3 kcal/mol stronger than that of the H···O bond. This estimation for the bonding energy of the bifurcated H-bond is further confirmed by the stabilization energies in partially optimized base pairs **d6** (11.78 kcal/mol for nonplanar motifs and 12.11 kcal/mol for planar motifs). It should be noted that the bifurcated H-bond discussed above involves the protons that originate from the single base. In many tetrads, two protons in the bifurcated H-bond are not from the same base as can be seen from Figure 5 (**t3** and **t4**). To estimate the bonding energy of the bifurcated H-bond involving three bases, an isoG trimer extracted from the tetrad conformer **t4** was optimized at the B3LYP/6-311G(d,p) level of theory (Figure 4). The stabilization energy of this trimer amounts to 39.85 kcal/mol after both ZPE and BSSE corrections.[34] Notice that this trimer comprises one normal O···H bond, two N···H bonds, and one bifurcated H-bond which contains protons from different bases. The bonding energy assigned to this bifurcated H-bond is 19.85 kcal/mol according to the previous bonding energy assignment for the O···H and N···H bonds, approximately two times as that of the single O···H bond. In this way, the bifurcated H-bond can be viewed

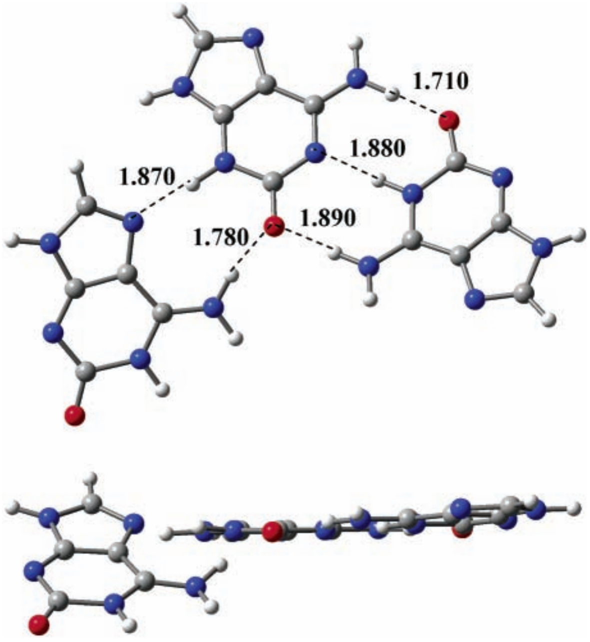


**Figure 4.** Two different views of the isoG trimer.[30] This trimer comprises one normal O···H bond, two N···H bonds, and one bifurcated H-bond which contains protons from different bases. The bonding energy assigned to this bifurcated H-bond is 20.85 kcal/mol, approximately two times that of the single O···H bond. In this way, the bifurcated H-bond can be viewed as two individual O···H bonds where two protons originate from two bases.

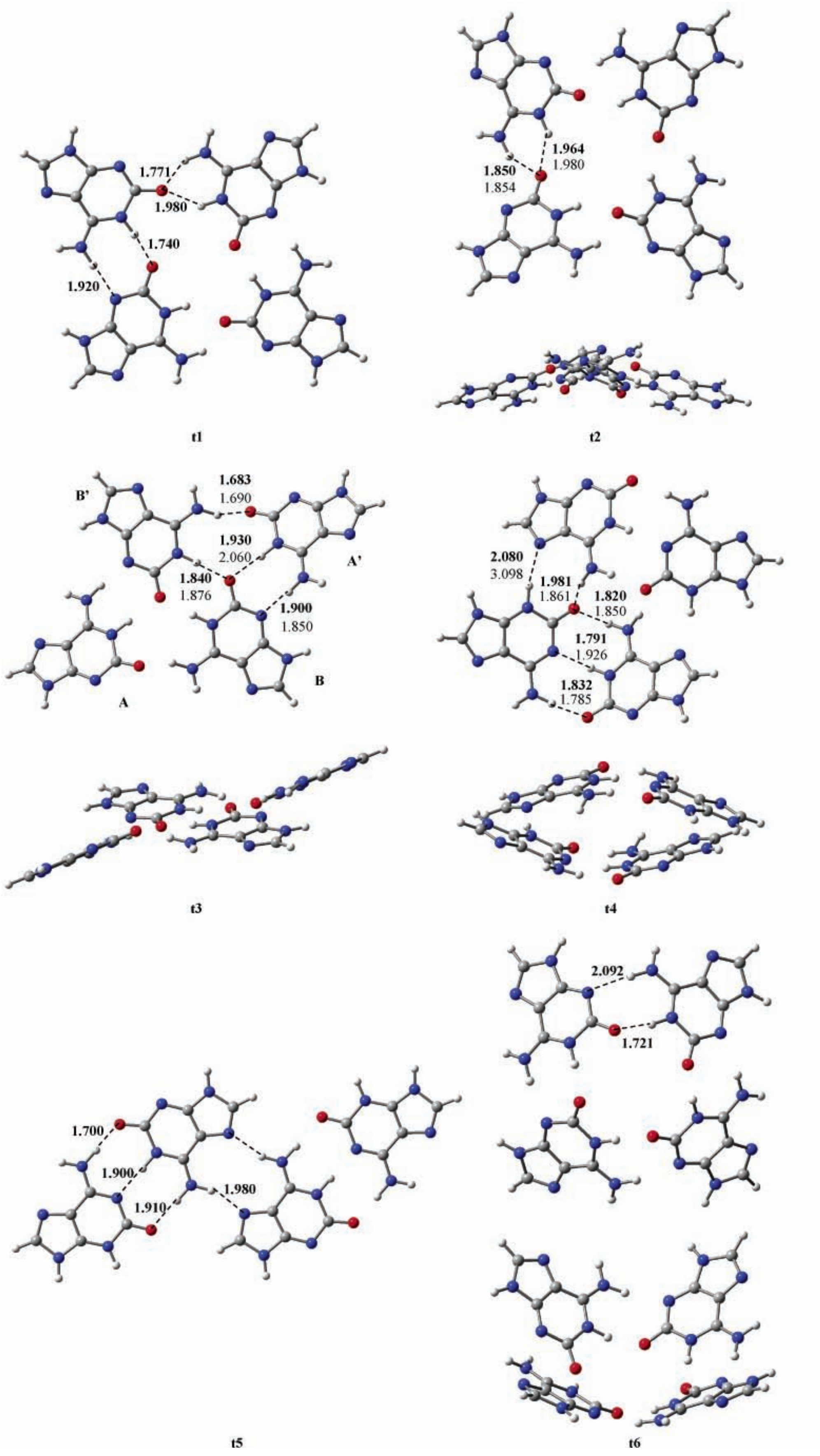


**Figure 5.** Geometries of different conformers of the isoG tetrad with side views. Parameters in bold: local minima structures; plain: planar form. **t6** is not the fully optimized structure.

**TABLE 2: Energy Properties of isoG Tetrads Evaluated at the B3LYP/6-311G(d,p) Level**[a]

| | symmetry | composition | $E$/hartree | $\Delta E$/kcal mol$^{-1}$ | $E^\circ$/hartree | $\Delta E^\circ$/kcal mol$^{-1}$ | $\Delta E^\circ$ (BSSE) | $\nu$ cm$^{-1}$ |
|---|---|---|---|---|---|---|---|---|
| tetrad | | | | | | | | |
| **t1** | $C_{2h}$ | 1H-isoG | −2170.902859 | −87.85 | −2170.431022 | −83.18 | −74.22 | 5 |
| **t2** | $S_4$ | 1H-isoG | −2170.898653 | −85.21 | −2170.426373 | −80.26 | −71.86 | 5 |
| | $C_4$ | | −2170.897675 | −84.59 | −2170.426057 | −80.06 | −71.66 | |
| | $C_{4h}$ | | −2170.897671 | −84.59 | −2170.426138 | −80.11 | −71.71 | −6 |
| **t3** | $C_i$ | 1H-isoG | −2170.898803 | −85.30 | −2170.426742 | −80.49 | −72.09 | 1 |
| | $C_{2h}$ | | −2170.897007 | −84.17 | −2170.425800 | −79.90 | −71.06 | |
| **t4** | $C_1$ | 1,3H-isoG | −2170.889990 | −80.43 | −2170.416843 | −74.25 | −66.99 | 5 |
| | $C_i$ | | −2170.889903 | −80.37 | −2170.416743 | −74.19 | −66.93 | −6 |
| | $C_{2h}$ | | −2170.884583 | −77.03 | −2170.412844 | −71.74 | −64.58 | |
| **t5** | $C_2$ | 1,3H-isoG | −2170.878846 | −73.43 | −2170.407730 | −68.53 | −61.27 | 5 |
| **t6** | $C_4$ | 1H-isoG | −2170.892253 | −81.85 | −2170.42076 | −76.74 | −68.34 | |
| monomer | | | | | | | | |
| 1H-isoG | | | −542.6907167 | | −542.574618 | | 2.10* | |
| 3H-isoG | | | −542.6901951 | | −542.574640 | | 1.53* | |

[a] $E^0$ and $\Delta E^0$ are vibrational zero-point energy corrected, where $\Delta E = E(\text{tetrad}) - 2E(\text{monomer1}) - 2E(\text{monomer2})$, $\Delta E^0 = E^0(\text{tetrad}) - 2E^0(\text{monomer1}) - 2E^0(\text{monomer2})$, and $\Delta E^0(\text{bsse}) = E^0(\text{tetrad}) - 2E^0(\text{monomer1}) - 2E^0(\text{monomer2}) + 2\text{BSSE}(\text{monomer1}) + 2\text{BSSE}(\text{monomer2})$. The asterisk is the BSSE of monomer. $\nu$ is the smallest vibrational frequency for the optimized local minima and the imaginary frequency for transition state structures.

as two individual O···H bonds when two protons originate from different bases. We shall follow this viewpoint in the discussion hereafter.

The electrostatic potential maps of the isoG base pairs are displayed in Figure 3. It is clear from the electrostatic potential map that the nonplanar conformation of the base pairs arises from the local electrostatic repulsion between the bases. Specifically, a positive−positive repulsion between the charge located around H8 and H(N9)′ in **d2** and around H8 and H(N6)′ in **d3** forces these base pairs away from the plane. Both the positive-positive repulsion around H(N6) and H(N9)′ and the negative-negative repulsion around O2 and O2′ in **d4** twist this pair. There is no obvious local electrostatic repulsion in **d1** and **d5**. This is consistent with their planar or near planar geometric characteristics.

**Tetrads of isoG.** There are five different forms of the isoG tetrad located as local minima on the potential energy surfaces through full geometry optimization (Figure 5) of which two conformers present the planar structure (**t1** and **t5**) as was confirmed by subsequent analytical vibrational frequency calculations.

The most stable **t1** conformer of the isoG tetrad is held together through four pairs of normal and two pairs of bifurcated H-bonds. If we consider one pair of bifurcated H-bond as two O···H bonds, this system possesses eight hydrogen bonds. It should be noted that this conformer had not been yet proposed in the literature. Base pair **d4** can be easily recognized in this tetrad. Two **d4**s are linked together through bifurcated H-bonds between O2 of one isoG and H(N6)′ and H1′ of the other to form a stable cyclic planar structure with $C_{2h}$ symmetry. The atomic distances in the bifurcated H-bonds amount to 1.771 Å for O2···H(N6)′ and 1.980 Å for O2···H1′, respectively. It is important to note that the H-bond length of H(N6)···N3′ and H1···O2′ in **t1** has been reduced to 1.740 and 1.920 Å, respectively, about 0.22 Å shorter than those corresponding bonds in **d4**. In this cyclic tetrad conformer, each isoG provides two protons to its neighbor and at the same time receives two protons from the other neighbor to form the H-bonds. In this way, the unbalanced charge that appears in **d4** is compensated in **t1** by the formation of a cyclic motif through bifurcated H-bonding.

The **t2** tetrad with $S_4$ symmetry is totally stabilized by the bifurcated H-bond pattern that resembles the pattern of the normal guanine tetrad.[8,9] The $C_4$ symmetry conformer characterized by the same bonding pattern has been predicted to be nearly planar. However, a subsequent vibrational frequency analysis shows that it is neither a local minimum nor a transition state structure. The related transition state structure possesses $C_{4h}$ symmetry with the same bifurcated H-bonds. It is the ability to form bifurcated H-bonding that provides the possibility for isoGs to adopt this near planar structure. The small base-pairing sector angle between the hydrogen donor and acceptor faces in isoG (67°)[1] prevents it from forming a near planar structure in $C_4$ symmetry through the Hoogsteen bases-pair motif. The O2···H1′ bond length is calculated to be 1.960 Å in **t2**. These bond distances are the same in all forms with different symmetry. A slightly shorter bond distance has been detected for the other branch of the bifurcated H-bond. In the $S_4$ symmetry conformer, the O2···H(N6)′ distance amounts to 1.853 Å, 0.03 Å less than that at the $C_4$ and $C_{4h}$ symmetry species. As a comparison, the corresponding atomic distances in the G tetrad at the same theoretical level have been reported to be 2.063 and 1.945 Å, respectively.[8]

The **t3** conformer suggested earlier by Golas et al.[7] has been predicted to be less stable than **t1**, though they are similar in molecular geometries as can be seen from Figure 5. Instead of bifurcated H bonds, two **d4** isoG dimers are held through O2···H(N6)′ and O2″···H1′ bonding. The H-bonding motif in the **d5** base pair can be identified in **t3** (B, B′ in Figure 5) in which, however, one branch of the bifurcated H-bonding (O2···H(N6)′) has been replaced by the O2 atom of the isoG(A, A′) on the other side. The H(N6)···N3′ bond distance has been evaluated to be 1.900 Å in **t3** which is close to that in **t1**. The H1···O2′ bond length in **t3** has been predicted to be 1.930 Å, 0.03 Å shorter than that in **d4** and 0.20 Å longer than that in **t1**. Each isoG(A) in **t3** receives one proton and donates two protons through H-bonding, whereas each isoG(B) accepts three and contributes two protons. Therefore, an unbalanced charge distribution is expected for this tetrad.

The **t4** and **t5** conformers consist of both 1H-isoG and 3H-isoG. Like the discussed similarity between the **d1** base pair and the GC base pair, the planar form of the **t4** isoG tetrad resembles the GCGC tetrad. The length of the H-bonds which hold the two **d1** base pairs amounts to 1.860 Å, whereas the length of the H-bonds which hold the two GC pairs in the GCGC tetrad is 1.989 Å at the same level of theory.[19] **t4** is then expected to be more stable than the GCGC tetrad. However, the energy analysis below shows that this is not confirmed by

**TABLE 3: Hydrogen Bond Energy Estimation Based on the IsoG Bases Pairs and the Bonding Energy Predictions for IsoG Tetrads Based on These H-bond Energies**[a]

| bond type | H...O | H...N | bifurcated | BE (estimated) | $-\Delta E°$ (BSSE) |
|---|---|---|---|---|---|
| BE/kcal $\text{mol}^{-1}$ | 10 | 5 | 13 | | |
| | bond number contained | | | | |
| tetrad | | | | | |
| **t1** | 2 | 2 | 2 | 56 | 74.22 |
| **t2** | | | 4 | 52 | 71.86 |
| **t3** | 6 | 2 | | 70 | 72.09 |
| **t4** | 6 | 2 | | 70 | 66.93 |
| **t5** | 4 | 4 | | 60 | 61.27 |
| **t6** | 4 | 4 | | 60 | 68.34 |

[a] BE is the bond energy in kcal/mol.

calculations. The vibrational frequency calculation reveals that the planar form is not the local minimum on the potential energy surface. The local minimum structure of **t4** is formed from the two isoG base pairs of the **d1** bend in the opposite directions so that the N7 of 1H-isoG in one of the **d1** pairs can accept the H3 proton from the other **d1** to create another H-bond. As a result, the number of H-bonds in **t4** is as large as ten. However, the longer H-bond lengths predicted for H(N6)•••O2′ (1.820 Å) and for N1•••H3′ (1.970 Å) compared to those in **d1** suggest that the bending between the pairs weakens the H-bonding.

Two base pairs of **d1** are bonded through the H-bonding pattern of **d3** to form the planar tetrad structure **t5**. **t5** is the only isoG tetrad not adopting the cyclic form. The H-bonds between 1H-isoG and 3H-isoG do not change much compared to those in **d1**. The elongation by 0.03 Å observed for O2•••H(N6)′ is mainly due to the fact that in the tetrad the other proton attached to the N6 of 1H-isoG also forms an H-bond with the N7 of the opposite 1H-isoG.

Because of the small base-pairing sector angle between the hydrogen donor and acceptor planes in isoG, the bowl-shaped structure is the only possibility to build up an isoG tetrad through the Hoogsteen bases-pair motif. However, we were unable to locate it as the local minimum by the full optimization technique. Applying the constrained optimization, a bowl-like **t6** conformer was obtained. However, **t6** is neither a local minimum nor a transition state structure as revealed by the frequency analysis. Without constraints, only the addition of a cation into the central area of the tetrad leads to a bowl-shaped structure.[4]

isoG tetrads are stable H-bonded complexes as revealed in this study. The stabilization energy has been evaluated to be as high as 70∼80 kcal/mol for all studied conformers. Except for **t5** which adopts a noncyclic pattern, all other tetrads are more stable than the normal guanine tetrad ($\Delta E = 76.80$ kcal/mol at the same theoretical level, see ref 13 as well as this study). The difference of the stabilization energy between the planar and nonplanar conformers is basically negligible for **t2** and **t3** (Table 2). Even for **t4** in which the planar structure has eight H-bonds, two less than the corresponding nonplanar conformer, the energy difference amounts to only 2.5 kcal/mol. As long as they can be stacked between layers of other tetrads, the isoG tetrad should easily adopt planar forms.

The **t1** tetrad is the most stable one. It is about 3 kcal/mol more stable than other conformers and about 11 kcal/mol more stable than the G tetrad. Applying the bond energy values derived for the base pairs above, stabilization energy of only 56 kcal/mol is estimated for **t1** (Table 3). Clearly, about 18 kcal/mol (after ZPE and BSSE corrections) should be attributed to the cooperative effects in the cyclic form tetrad.[17] Stabilization energy of **t2** amounts to 71.9 kcal/mol after corrections which is 20 kcal/mol larger than the sum of the bond energy of four isolated bifurcated H-bonds. The cooperative effects seem to be much stronger in the symmetric H-bonding donor−acceptor systems. The total binding energy of **t1** and **t2** is comparable to that found in isoG-quintet (without BSSE, they are 21.96 kcal/mol for **t1** and 21.30kcal/mol for **t2** vs 21.91 kcal/mol for isoG quintet[25]). A much smaller difference between the calculated stabilization energy ($\Delta E°(\text{BSSE}) = -71.1$ kcal/mol) and the estimated bond energy (70 kcal/mol) has been observed for **t3**, about 1 kcal/mol. The unbalanced number of proton donators and acceptors among the bases suggests that there should be little or no cooperative effects in this complex. Although the nonplanar local minimum conformer of **t4** has two H-bonds more than the planar form, the stabilization energy reward for these extra H-bonds is small (2.5 kcal/mol as discussed above). The bending of the base pairs in **t4** eliminates most of the energy gain from the extra H-bonding. It is important to notice that

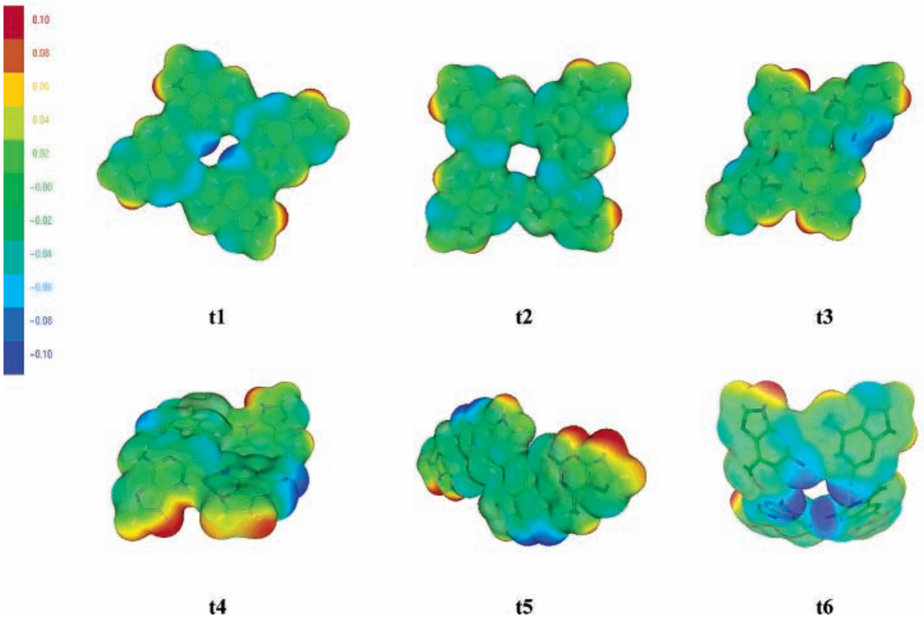


**Figure 6.** Electrostatic potential maps of the isoG tetrads. Electrostatic potential of isoG tetrads are mapped onto the surface of the electron density of 0.002 unit.

the planar form of **t4** shares a similar bonding pattern with the GCGC tetrad. However, the bonding energy of this planar **t4** is about 6.9 kcal/mol less than that of GCGC.[19] The bond energy estimation for this conformer is 70 kcal/mol, about 5.5 kcal/mol larger than the calculated value of 64.5 kcal/mol in this study. A detailed analysis of the bonding pattern of this planar conformer enables us to identify that 3H-isoGs accept three protons and donate only one in H-bonding, whereas 1H-isoGs contribute three and accept one. This highly unbalanced proton exchange is expected to destabilize the complex. On the other hand, each cytosine and guanine in the GCGC tetrad accepts two and at the same time donates two protons to the bonding. The cooperative effect is then expected for the GCGC tetrads. Because **t5** is the only tetrad not in a cyclic form, cooperative effects are not expected for this conformer. The bond energy estimation of 60 kcal/mol seems to be a good approximation for the stabilization energy of 61.3 kcal/mol in **t5**. Although **t6** is not the local minimum on the potential energy surface, a high stabilization energy has been evaluated (81.85 kcal/mol without corrections) which is about 5 kcal/mol larger than that in the G tetrad. The existence of cooperative effects is also obvious in **t6**. The corrected stabilization energy amounts to 68.3 kcal/mol, about 8.3 kcal/mol larger than the 60 kcal/mol bond energy estimation. The extraordinary stability of the isoG tetrad observed in experiments has been suggested to be the consequence of additional base-sugar H-bonds.[4] Our results show that the high stabilization energy of the isoG tetrad itself also plays an important role.

In their recent study on the isoG polyads, Meyer and Suhnel concluded that the most stable hydrogen bonded isoG tetrad possesses the $C_{4h}$ symmetry.[24] However, our results obtained using a larger bases set indicate that this conformer is not a local minimum on the PES. Rather, it is a transition state structure between two symmetrical local minima with $S_4$ symmetry, which resembles the G tetrad.[14] Even this $S_4$ symmetry conformer is not the most stable hydrogen bonded isoG tetrad, but the most stable conformer of IsoG tetrad is **t1** ($C_{2h}$ symmetry).

An electrostatic potential map offers a convenient way to predict the possible metal interaction sites for biological systems. The ESP maps depicted in Figure 6 exclude the location of metal ions in the central area of the **t3**, **t4**, and **t5** conformers. The ESP map of **t4** suggests a possible cation interaction site around O2 and N3 of 1H-isoG. The un-symmetric negative electrostatic potential center of **t1** and **t2** should not be favored by cations interacting mainly through electrostatic attraction. In fact, the flat potential energy surface around **t2** ($S_4$ symmetry) enables its conversion into $C_4$ symmetry under the influences of an approaching cation and therefore into the **t6** conformer which is clearly a preferential form to be a cation host as can be seen from its ESP map.

## Conclusions

An analysis of the energetic properties of isoG base pairs enables us to assign the hydrogen bond energy to isoG−iosG interactions and to apply it in the study of the energy properties of isoG tetrads. The approximate estimation of the H-bond energy amounts to 5 kcal/mol for the H···N bond, 10 kcal/mol for the H···O, and 13 kcal/mol for the bifurcated H-bond in which two protons originate from one base.

The energy difference between the local minima considered here and their corresponding planar structures of the base pairs is insignificant. An easy flapping motion for the bases has also been confirmed by the vibrational mode analysis. The nonplanarity of the isoG base pairs arises from the electrostatic repulsion between the bases in the pair.

Five different isoG tetrads have been predicted to be the local minima on the potential energy surface. The relative stability is as follows: **t1** > **t3** > **t2** > **t4** > **t5**. A high stabilization energy has been revealed for these isoG tetrads. All cyclic isoG tetrads are more stable than guanine tetrads.

The difference in stabilization energy between the planar and nonplanar conformers is basically negligible. As long as they can be stacked between layers of other tetrads, isoG tetrads should easily adopt a planar form. Introducing a cation, especially alkali metal ions, in the central area of isoG tetrads leads inevitably to a bowl-like structure as revealed by the electrostatic potential maps.

Applying the bond energy assignment, we are able to investigate the cooperative effect of hydrogen bonding in the cyclic form of the tetrads. The cooperative effect of hydrogen bonding is crucial for highly stabilized isoG tetrads. The balanced transfer of proton among all componenets of the isoG tetrad is necessary to build up the cooperative effect. Bifurcated H-bondings are important in the formation of isoG tetrads. This type of bonding pattern provides alternatives for isoG to form different stable tetrad conformers which have the potential to develop diversity in biological systems.

**Acknowledgment.** This research project in the PRC was supported by the “Knowledge Innovation Program” and the “Introducing Outstanding Overseas Scientists Project,” Chinese Academy of Sciences. In the U.S.A. the project was supported by NIH Grant No. G1 2RR13459-21 and NSF Grant No. 9805465.

## References and Notes

(1) Chaput, J. C.; Switzer, C. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10614−10619.
(2) Eschenmoser, A.; Loewenthal, B. *Chem. Soc. Rev*. **1992**, *21*, 1−16.
(3) Seela, F.; Kroschel, R. *Bioconjugate. Chem.* **2001**, *12*, 1043−1050.
(4) Tirumala, S.; Davis, J. T. *J. Am. Chem. Soc*. **1997**, 119, 2769−2776.
(5) Seela, F.; Wei, C.; Melenewski, A. *Nucleic Acids Res*. **1996**, *24*, 4940−4945.
(6) Seela, F.; Wei, C. F. *Chem. Commun*. **1997**, 1869−1870.
(7) Golas, T.; Fikus, M.; Kazimierczuk, Z.; Shugar, D. *Eur. J. Biochem.* **1976**, *65*, 183−192.
(8) Gu, J.; Leszczynski, J. *J. Phys. Chem. A* **2000**, *104*, 6308−6313.
(9) Gu, J.; Leszczynski, J.; Bansal, M. *Chem. Phys. Lett.* **1999**, *311*, 209−214.
(10) Davis, J. T.; Tirumala, S.; Jenssen, J. R.; Radler, E.; Fabris, D. *J. Org. Chem*. **1995**, *60*, 4167−4176.
(11) Krishnamurthy, R.; Pitsch, S.; Miniton, M.; Miculka, C.; Windhab, N.; Eschenmoser, A. *Angew. Chem*. **1996**, *108*, 1619−1623.
(12) Cai, M.; Marlow, A. L.; Fettinger, M. J.; Fabris, D.; Haverlock, T. J.; Moyer, B. A.; Davis, J. T. *Angew. Chem*. **2000**, *112*, 1339−1341.
(13) Malashkevich, V. N.; Kammerer, R. A.; Efimov, V. P.; Schulthess, T.; Engel, J. *Science* **1996**, *274*, 761−765.
(14) Meyer, M.; Brandl, M.; Suhnel, J. *J. Phys. Chem. A* **2001**, *105*, 8223−8225.
(15) Meyer, M.; Steinke, T.; Brandl, M.; Suhnel, J. *J. Comput. Chem*. **2001**, *22*, 109−124.
(16) Meyer, M.; Schneider, C.; Brandl, M.; Suhnel, J. *J. Phys. Chem. A* **2001**, *105*, 11560−11573.
(17) Gu, J.; Leszczynski, J. *J. Phys. Chem. A* **2002**, *106*, 529−532.
(18) Gu, J.; Leszczynski, J. *J. Phys. Chem. A* **2000**, *104*, 1898−1904.
(19) Gu, J.; Leszczynski, J. *J. Phys. Chem. A* **2000**, *104*, 7353−7358.
(20) Gu, J.; Leszczynski, J. *J. Phys. Chem. A* **2001**, *105*, 10366−10371.
(21) Gu, J.; Leszczynski, J. *Chem. Phys. Lett.* **2001**, *335*, 465−474.
(22) Gu, J.; Leszczynski, J. *Chem. Phys. Lett.* **2002**, *351*, 403−409.
(23) Roberts, C.; Chaput, J. C.; Switzer, C. *Chem. Biol*. **1997**, *4*, 899−908.
(24) Meyer, M.; Suhnel, J. *J. Phys. Chem. A* **2003**, *107*, 1025−1031.
(25) Gu, J.; Leszczynski, J. *J. Phys. Chem. B* **2003**, *107*, 6609−6613.
(26) Gu, J.; Leszczynski, J. *J. Phys. Chem. A* **2000**, *104*, 7353−7358
(27) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(28) Lee, C.; Yang W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.

(29) Miehlich, B.; Savin, A.; Stoll H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200−206.

(30) Hehre, W. J.; Radom, L.; Schleyer, P. R.; Pople, J. A. *Ab initio Molecular Orbital Theory*; Wiley: New York, 1986.

(31) Mebel, A. M.; Morokuma, K.; Lin, C. M. *J. Chem. Phys.* **1995**, *103*, 7414−7421.

(32) Johnson, B. G.; Gill, P. M. W.; Pople, J. A. *J. Chem. Phys.* **1993**, *98*, 5612−5626.

(33) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision D.3; Gaussian, Inc.: Pittsburgh, PA, 1998.

(34) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.

(35) Leszczynski, J. *Int. J. Quantum Chem. Quantum Biol. Symp.* **1992**, *19*, 43.