J. Phys. Chem. A **2001,** *105,* 8223−8225

Are Guanine Tetrads Stabilized by Bifurcated Hydrogen Bonds?

Michael Meyer,*[†] Maria Brandl,[‡] and Jürgen Sühnel[‡]

Revotar Biopharmaceuticals AG, Neuendorfstrasse 24b, D-16761 Hennigsdorf, Germany, and Institut für Molekulare Biotechnologie, Biocomputing Group, Beutenbergstrasse 11, D-07745 Jena, Germany

Received: March 28, 2001; In Final Form: July 2, 2001

We have carried out Hartree−Fock and B3LYP hybrid density functional studies with various basis sets to analyze the geometries and relative energies of different guanine tetrad conformations with Hoogsteen-type and bifurcated hydrogen bonds. B3LYP/DZVP calculations yield energy minima for nonplanar tetrad structures with both types of hydrogen bond patterns. At the highest B3LYP/6-311+G(2d,p)//B3LYP/6-311G(d,p) level of theory, a planar conformation with a bifurcated hydrogen bond pattern is the most stable structure we found, whereas at the B3LYP/6-311G(d,p) level, the nonplanar Hoogsteen tetrad is the more stable one. The energy differences between structures with Hoogsteen-type and bifurcated hydrogen bonds as well as between nonplanar S_4 symmetric and planar C_{4h} symmetric structures are very small and depend on the choice of the basis. Finally, HF/6-311G(d,p) calculations show a significantly different potential energy surface compared to the B3LYP method because the local energy minima for the Hoogsteen-type structures are missing.

Introduction

Guanine (G) tetraplexes,[1,2] four-stranded nucleic acid structures occurring in telomeric regions of chromosomes, and G-tetrads, structural subsets of these biopolymers, have attracted considerable interest in biochemistry and in computational chemistry. It is well-known that metal ions are essential in the formation of these tetraplex structures. In addition to experimental studies, molecular dynamics simulations provided information about interactions between the tetraplex and cations.[3] These calculations have been supplemented recently by quantum-chemical studies of G-tetrads at the Hartree−Fock (HF) and density functional (DFT) levels of theory.[4−6]

In general, DFT calculations yield base−base interaction energies and base pair structures with a good quality.[7,8] For the G-tetrad however, recent quantum-chemical studies have come to apparently contradicting results. Gu et al. have optimized the G-tetrad structures at the HF and B3LYP levels and concluded that the isolated tetrad without a metal ion is stabilized by bifurcated hydrogen bonds (H-bonds) formed by the hydrogen atoms H1 and H22 and the acceptor atom O6 (Figure 1).[4,5] According to these authors, the presence of a metal cation induces the normal Hoogsteen-type structure with two distinct N1−H1...O6 and N2−H22...N7 H-bonds (Figure 2). The latter H-bond pattern is also found in experimental NMR or X-ray structures where metal ions are required for tetrad formation. In contrast to the theoretical studies mentioned above, we have focused on the Hoogsteen-type base pairing and noticed no change in the H-bond pattern when the cations were removed from the complex structures.[6] Here we present additional calculations to shed light on the contradicting results obtained by previous quantum-chemical studies of G-tetrads.

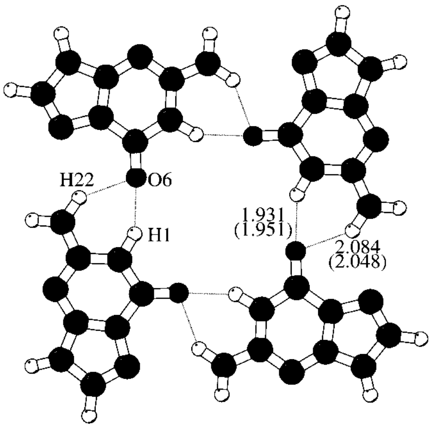

Figure 1. Structure of the G-tetrad with bifurcated N1−H1...O6 and N2−H22...O6 H-bonds. H-bond distances determined with the B3LYP/6-311G(d,p) method are given for S_4 and C_{4h} symmetry (C_{4h} data in parentheses). The Figures were generated with MOLSCRIPT.[23]

Methods

The structures have been optimized with GAUSSIAN98[9] using the B3LYP hybrid density functional method[10,11] with the 6-31G(d,p) and 6-311G(d,p) basis sets[12,13] and the DZVP basis set optimized for DFT calculations.[14] Energy minima have been verified with subsequent frequency calculations at the B3LYP/DZVP level. Subsequent single-point calculations have been carried out using the TZVP basis[15] and multiple polarized basis sets and diffuse functions[16,17] with the "Tight" SCF convergence criterion and the default grid for integration. Additionally, we carried out calculations using the UltraFineGrid and the "Tight" optimization criterion to check the numerical accuray. Complexation energies of the tetrad have been estimated using the counterpoise approach.

Results

The structures have been investigated at the B3LYP level for both H-bond types at planar C_{4h} and nonplanar C_4 and S_4

---

* Corresponding author. Phone: +49 3302 202 5061, Fax: +49 3302 202 5030. E-mail: m.meyer@revotar-ag.de.

[†] Revotar.

[‡] Institut für Molekulare Biotechnologie.

10.1021/jp011179i CCC: $20.00

TABLE 1: Total Energies E^a of G-tetrads for Different Symmetries and H-bond Patterns

	E (H)		E (H)	
	Hoogsteen		bifurcated	
method	C_{4h}	S_4	C_{4h}	S_4
HF/6-311G(d,p)	–	–	−2158.21951	**−2158.21953**
B3LYP/DZVP	−2170.63602	−2170.63641[b]	−2170.63678	**−2170.63692**[c]
B3LYP/TZVP//B3LYP/DZVP	−2171.00314	**−2071.00439**	−2171.00286	−2071.00397
B3LYP/6-31G(d,p)	−2170.38820	**−2170.38882**	−2170.38654	−2170.38687
B3LYP/6-311G(d,p)	−2170.91356	**−2170.91404**	−2170.91308	−2170.91329
B3LYP/6-311G(2d,p)//B3LYP/6-311G(d,p)	−2170.95626	−2170.95736	−2170.95726	**−2170.95815**
B3LYP/6-311G+G(d,p)//B3LYP/6-311G(d,p)	−2170.95935	**−2170.95966**	−2170.95949	−2170.95925
B3LYP/6-311G+G(2d,p)//B3LYP/6-311G(d,p)	−2171.00462	−2171.00521	**−2171.00595**	−2171.00552

[a] The lowest energy at each level is printed in bold. [b] Zero-point vibrational energy: 0.47104 kcal/mol.[6] [c] Zero-point vibrational energy: 0.47132 kcal/mol.

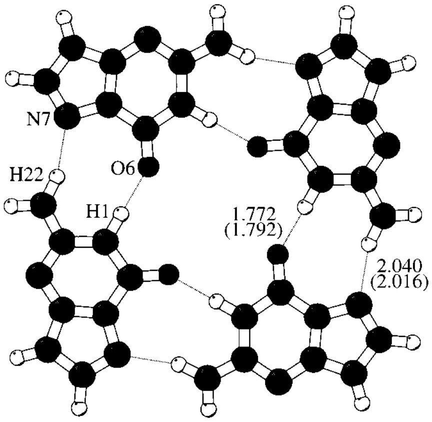

Figure 2. Structure of the Hoogsteen-type G-tetrad with N1−H1...O6 and N2−H22...N7 H-bonds. H-bond distances determined with the B3LYP/6-311G(d,p) method are given for S_4 and C_{4h} symmetry (C_{4h} data in parentheses).

TABLE 2: Energy Differences ΔE between the Hoogsteen-type and Bifurcated Structures of G-tetrads for Different Symmetries

	ΔE (kcal/mol)	
method[a]	C_{4h}	S_4
HF/6-311G(d,p)	b	b
B3LYP/DZVP	0.48	**0.32**
B3LYP/TZVP//B3LYP/DZVP	−0.17	**−0.26**
B3LYP/6-31G(d,p)	−1.04	**−1.22**
B3LYP/6-311G(d,p)	−0.30	**−0.47**
B3LYP/6-311G(2d,p)//B3LYP/6-311G(d,p)	−0.51	**−0.50**
B3LYP/6-311+G(d,p)//B3LYP/6-311G(d,p)	0.09	**−0.26**
B3LYP/6-311+G(2d,p)//B3LYP/6-311G(d,p)	**0.83**	0.19

[a] Bold letters indicate the symmetry of the most stable structure at each level. [b] No energy minimum for the Hoogsteen-type structure was found at this level.

symmetries, respectively. Nonplanar structures can arise from the pyramidal amino group of G. Both C_{4h} symmetric structures show four imaginary frequencies and thus do not correspond to local energy minima at the B3LYP/DZVP level. For C_4 symmetry, we found a structure of the bifurcated conformer that corresponds to a local energy minimum, even though it is very similar to the C_{4h} symmetry one (root-mean-square deviation of the atomic coordinates 0.12 Å, energy difference 0.02 kcal/mol). Since the differences of geometries and energies are small, we focus on the C_{4h} and S_4 symmetric structures for both H-bond patterns.

With all basis sets used for optimization, DZVP, 6-31G(d,p) and 6-311G(d,p), the nonplanar S_4 symmetric structures are somewhat more stable than the planar C_{4h} symmetric ones (Table 1). In the former structures, the amino group nitrogen atoms are slightly pyramidal, and the attached hydrogen atoms are located above and below the base planes in an alternating manner. For both H-bond patterns, the change of the symmetry from C_{4h} to S_4 is accompanied by an elongation of the H1...O6 distance and a shortening of the distance between H22 and the acceptor atoms O6 and N7, respectively (Figures 1 and 2). At the B3LYP/DZVP level, the bifurcated structure is 0.32 kcal/mol more stable, whereas according to the energies obtained with both Pople-type basis sets mentioned above, the Hoogsteen-type structure is more stable (Table 2). The energy difference between structures with either bifurcated or Hoogsteen-type H-bonds is largest at the B3LYP/6-31G(d,p) level (−1.11 kcal/mol), and it decreases slightly in magnitude to −0.26 kcal/mol when larger basis sets up to 6-311+G(d,p) are used for single-point calculations. For such small energy differences, even the −0.18 kcal/mol difference for the zero-point vibrational energies begins to become important.

At the B3LYP/6-311+G(2d,p)//B3LYP/6-311G(d,p) level, a remarkable change in the relative energies occurs. The bifurcated complex is more stable than the Hoogsteen tetrad by 0.83 kcal/mol at C_{4h} symmetry and 0.19 kcal/mol at S_4 symmetry (Table 2). Hence, in contrast to all other calculations at lower levels, the planar bifurcated structure is the most stable one. However, the energy difference relative to the S_4 symmetric Hoogsteen-type structure is only 0.46 kcal/mol.

The optimization of structures based on a higher grid for integration and the "tight"convergence criterion indicate that the total energies are stable within 0.04 mH and changes of the relative energies between different conformations are on the order of 0.01 kcal/mol at the B3LYP/6-311G(d,p) level.

Another surprising result was obtained from optimizations at the HF level. Starting with the Hoogsteen structure, the bases move relative to each other during optimization so that a tetrad stabilized by bifurcated H-bonds was obtained for both symmetries. Therefore, probably no local HF energy minima exists for the Hoogsteen-type complex. Thus, the HF energy surface seems to be quite different from the ones obtained with the BLYP[6] and B3LYP approaches.

Discussion

Our calculations show that the potential energy surface of the G-tetrad depends significantly on the computational approach. At the HF/6-311G(d,p) level, the energies of the bifurcated tetrad structures at C_{4h} and S_4 symmetry are very similar to the ones reported previously for C_4 symmetry.[4,22] The missing energy minimum of the Hoogsteen-type complex at the HF/6-311G(d,p) level as compared to the BLYP[6] and B3LYP calculations indicates a potential risk in structure calculations of complex molecules. To reduce the computational effort, structures may be optimized first at a lower level to generate a good starting point for subsequent optimizations with more accurate and computationally demanding methods. When both potential energy surfaces have corresponding energy minima this is an efficient strategy, however, it may be misleading if certain energy minima are only present at a higher or lower level of theory.

The question of accuracy arises directly from the different results obtained by both methods. Variations of the SCF convergence criterion did not lead to significant changes of the calculated energies. Several comparative studies indicate that the economical B3LYP method is able to provide similar results as the MP2 method or single-point MP2 calculations based on structures optimized with the HF method.[8,18−21] To our knowledge, comparisons at higher levels are not available for base complexes.

As indicated by the data of Table 1, the choice of the basis set is also crucial for the determination of the relative energies and the structures. Characteristic for S_4, and probably also C_4 symmetry structures, are pyramidal amino groups, whereas the amino groups are planar at C_{4h} symmetry. Therefore, the dependence of the amino group inversion barrier on the method is likely to be an important factor. Šponer and Hobza reported an energy difference of 1.12 kcal/mol between the planar and the more stable pyramidal amino groups for G with the MP2/6-311G(2df,p) method.[8] This approach is currently not feasible for G-tetrads. B3LYP/6-311G(d,p) calculations predict an energy difference of only 0.56 kcal/mol. Therefore, methods including electron calculation at a higher level and larger basis might favor nonplanar conformations.

The present DFT calculations indicate that the energy differences between the Hoogsteen-type structure and the one with bifurcated H-bonds are very small. The B3LYP/6-311G+G(2d,p)//B3LYP/6-311G(d,p) approach shows that the bifurcated H-bond pattern leads to a slightly more stable complex geometry at C_{4h} symmetry and lower levels of theory favor the S_4 symmetry Hoogsteen structure. The latter conformation is also more stable than the one proposed previously for the bifurcated complex structure at the B3LYP/6-311G(d,p) level.[4,5,22] The BSSE corrected complexation energy −75.07 kcal/mol for the planar bifurcated tetrad structure exceeds the one of the Hoogsteen type structure (−73.77 kcal/mol). These energies determined at the highest single-point level are only somewhat smaller than the results calculated previously with smaller basis sets.[6]

Conclusions

B3LYP hybrid density functional calculations using various basis sets indicate that G-tetrads can adopt two conformations. One of them, also obtained at the HF level, is linked by bifurcated H-bonds, whereas the other Hoogsteen-type structure was found only at the DFT level. The preference for planar or nonplanar structures with the bifurcated or Hoogsteen-type H-bond pattern depends on the basis set. For example, at the B3LYP/6-311G(d,p) level, the nonplanar Hoogsteen-type structure is the most stable one, whereas at the higher B3LYP/6-311G+G(2d,p)//B3LYP/6-311G(d,p) level, a planar structure with bifurcated H-bonds is preferred. The energy differences between the various conformations are very small, however. Therefore, it is not unlikely that investigations with full structure optimizations at higher levels of theory may lead to novel aspects. As the amino group barrier to inversion determined with multiple polarized basis sets at the MP2 level[8] exceeds the one from B3LYP calculations, nonplanar complexes of C_4 or S_4 symmetry may turn out to be more stable than the preferred planar structure found in this study. In view of the small energy differences and the basis set dependence of the results, we think that the type of the base−base interaction in G-tetrads is still open to discussion. The small energy differences between the conformations suggest further that forces exerted by the nucleic acid environment, ions and solvent can easily influence the H-bond pattern and the planarity or nonplanarity.

References and Notes

(1) Williamson, J. R. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, *23*, 703.

(2) Han, H.; Hurley, L. H. *Trends Pharm. Sci.* **2000**, *21*, 136.

(3) Špackovà, N.; Berger, I.; Šponer, J. *J. Am. Chem. Soc.* **1999**, *121*, 5519.

(4) Gu, J.; Leszczynski, J.; Bansal, M. *Chem. Phys. Lett.* **1999**, 311, 209.

(5) Gu, J.; Leszczynski, J. *J. Phys. Chem. A* **2000**, *104*, 6308.

(6) Meyer, M.; Steinke, T.; Brandl, M.; Sühnel, J. *J. Comput. Chem.* **2001**, *22*, 109.

(7) Fonseca Guerra, C.; Bickelhaupt, F. M.; Snijders, J. G., Baerends, E. J. *J. Am. Chem. Soc.* **2000**, 122, 4117.

(8) Hobza, P.; Šponer, *J. Chem. Rev.* **1999**, *99*, 3247.

(9) 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.; Stratman, 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.; Mennuci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokua, K.; Malik, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T. A.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzales, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzales, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *GAUSSIAN 98*; Gaussian, Inc.: Pittsburgh, PA, 1998.

(10) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(11) Lee, C.; Yang, G.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(12) Ditchfield, R.; Hehre, W. J.; Pople, J. A *J. Chem. Phys.* 1972, *56*, 2257.

(13) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.

(14) Godbout, E.; Salahub, D. R.; Andzelm, J.; Wimmer, E. *Can. J. Chem.* **1992**, *70*, 560.

(15) Sosa, C.; Andzelm, J.; Elkin, B. C.; Wimmer, E.; Dobbs, K. D.; Dixon, D. A. *J. Phys. Chem.* **1992**, 96, 6630.

(16) Clark, T.; Chandrasekar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, 4, 294.

(17) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265.

(18) Meyer, M.; Sühnel, J. *J. Biomol. Struct. Dyn.* **1997**, 15, 619.

(19) Šponer, J.; Burda, J. V.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn*. **1999**, *17*, 61.

(20) Brandl, M.; Meyer M.; Sühnel, J. *J. Am. Chem. Soc*. **1999**, *121*, 2605.

(21) Brandl, M.; Meyer, M.; Sühnel, J. *J. Phys. Chem. A* **2000**, *104*, 11177.

(22) Gu, J.; Leszczynski, J. *J. Phys. Chem. A* **2000**, *104*, 1898.

(23) Kraulis, P. *J. Appl. Crystallogr.* **1991**, *24*, 946.