Structure and Stability of DNA Base Trimers: An Electrostatic Approach

Savita S. Pundlik and Shridhar R. Gadre*

Department of Chemistry, University of Pune, Pune-411007, India

Received: July 28, 1997[⊗]

Triplexes involving major groove binding of a third oligomer to the DNA duplex structure have gathered much interest in recent years. The study of base trimer interactions in the gas phase is expected to provide useful information regarding orientational preferences and inherent stabilities. A recently developed electrostatic potential for intermolecular complexation (EPIC) model has been found to be quite useful for exploring the structures and energetics in base pairs [Gadre, S. R.; Pundlik, S. S. *J. Phys. Chem.* **1996**, *101*, 3298]. This model makes use of complementary electrostatic features of the interacting species that are determined by ab initio theory. We report here the investigations on various trimers including TAT, TAG, ATG, CGG, and TCG using this model. The overall good agreement of the trimer interaction energies with the corresponding single-point SCF values made at the model-predicted geometries reveals the suitability of the EPIC model for studying DNA base complexes.

1. Introduction

DNA triplexes have been found to be associated with certain detrimental effects including inhibition of transcription[1] and diseases caused by gene mutation.[2] These have attracted considerable attention recently owing to their probable use in treating genetic disorders via specifically restricting the activity of a particular gene sequence. For therapeutic purposes, it is possible to target the specificity of DNA through a third strand which can penetrate the cell wall and persist therein for hours.[3] The phenomenon of triplex formation, therefore, demands close scrutiny since there lies hope for curing gene-related diseases and possibly even controlling other biological activities.

The necessity for a detailed examination of DNA triplexes can thus hardly be overemphasized. This kind of a study mainly deals with the structures and stabilities of weak complexes involving bases such as adenine (A), thymine (T), guanine (G), and cytosine (C). The experimental[4,5] as well as theoretical[1] methods involved in such investigations generally aim at obtaining the information with a suitable incorporation of various external factors, so as to simulate the physiological conditions. Different experimental techniques have been developed for explaining the relative stabilities[3] and also the orientations attained by the bases during triplex formation. For example, it has been well-established now[5a,b,c] that TAT is one of the very stable triplexes with a Watson−Crick (WC) TA unit binding with another T in a Hoogsteen (Ht) manner. TAT and ATG have been found to have a single UV melting temperature[4e] in a 32-base oligonucleotide. This implies that the triplet oligomer almost concurrently separates into three constituent DNA strands at the melting point rather than first dissociating into a dimer and a monomer with subsequent dissociation of the dimer occurring at a higher temperature. NMR spectroscopic data have been used for finding which amino proton in guanine forms a hydrogen bond with thymine in an unusual structure of ATG.[4e] The CXG (X = A, T, G, or C) triplet repeat sequences have been studied with NMR and UV spectroscopic methods[2] in order to understand their structures and conformations. The chain length expansions of these have been found to be associated with neurodegenerative diseases.[2] One of the limitations in the experimental studies on oligomeric triplexes is that the bases in the third strand can specifically bind to either purine(Pu)-rich or pyrimidine(Py)-rich segments.[4c] In a prototypical NMR study[4c] involving TCG, a need has been felt to look into the details at a molecular level for understanding the recognition of a CG inverted site in an otherwise homo-Pu−Py duplex so as to build a general model for CG inversion in hetero-Pu−Py duplexes.

Theoretical methods such as molecular dynamics simulation are often used[4c,d] for making a comparison with the respective experimental findings. Mergny et al.[5e] have demonstrated that the instability effects of a mismatch nucleotide in a Ht-bonded strand of a triplex using a thermal dissociation experiment agree well with the results of molecular modeling studies.

As regards DNA base pairs, these are being constantly explored since the discovery of double-helical structure about fifty years ago.[6] On the theoretical front, enormous effort has been devoted for a better understanding[7] of base pair formation as well as exploring issues such as correlation effects,[7a] pyramidalization of the amino group,[7b] protonation, and tautomerism.[7c] Also, there have been continuous refinements in various empirical and semiempirical methods[7d] for making correct predictions regarding energetics in base pairs. Such methods, however, require time-consuming preparametrization, and yet there are reports questioning their reliability.[7a,8] For instance, the H-bond energy is not predicted correctly with the MM2 method.[8a] A given set of empirical parameters is certainly not universal. For example, the CO···NH bonds in GG and TT differ greatly (by about 50%) in stability.[8c] The structure of guanine is not predicted correctly with empirical potentials[8d] since attractive interactions with hydrogens of pyramidal amino groups are not taken into account. On the other hand, with the advent of better computational facilities, it has even been possible, albeit with a massive effort, to optimize the geometries of base pairs at a high level of ab initio theory.[7a] The electron correlation effects are not found to be significant in such H-bonded dimers. Thus, very accurate data on structures and energies of base pairs are now becoming available in the form of such "milestone" calculations.

The DNA trimers still await such rigorous characterization, although theoretical gas-phase stabilities at an empirical level have been reported[9] thirty years ago. A comparative study of stabilities in cyclic homo- and open heterotriplets involving inosine (I), uracil (U), A, G, and C is performed by evaluating

[⊗] Abstract published in *Advance ACS Abstracts,* October 15, 1997.

S1089-5647(97)02491-7 CCC: $14.00

a few energy component terms.[9] The cyclic form has been shown to be more energetically advantageous for I than the open one. Also, repulsion is introduced by interaction of the terminal U or I bases in the open structures[9] A + 2U and A + 2I. The interaction energy for open triplet CGC^+ is very large compared to other triplets. These studies have apparently not been pursued further for a variety of combinations of bases in a triplet which are now found to be important. On the other hand, a demand exists for understanding the molecular basis for triplet formation.[4c] It will be beneficial, therefore, to examine the inherent electronic interactions in gas-phase base trimers. Focusing on the built-in capacity of recognition at the molecular level would lead to newer insights, so that it may become possible to separate out the extent of effects caused by surrounding environment on the triplex stability. The bases, after pairing, still possess lone pairs as well as proton sites which may be accessed by yet another base, in a complementary manner, leading to the formation of a weakly bound trimer. Thus, formation of a base trimer essentially involves H-bond interactions which are known to be dominated by electrostatic components.[7a,10] High-level ab initio treatment, although very reliable, is extremely time consuming even for DNA base pairs with one of the fastest contemporary computer systems.[7a] It is felt that the use of the newly developed electrostatic potential for intermolecular complexation (EPIC)[11] model would be highly suitable for such studies because it is simple in application and retains the important ab initio characteristics of the species involved in an interaction. The appropriateness of this model has recently been demonstrated[12] for exploring DNA base pair interactions. In this paper, we examine DNA base trimers with the aid of the EPIC model.

2. Method

The importance of molecular electrostatic potential (MESP)[13] in molecular recognition phenomena has been time-honored now, leading to its incorporation into various models for weak binary complexes.[14] The strength of the EPIC model lies in the exploitation of the topography of MESP[15] for initial positioning of the two species, use of their complete electrostatic information, and almost no requirement of parametrization. To arrive at an optimum geometrical configuration for a binary complex (with rigid monomer geometries), one requires the molecular wave function and a suitable set of potential-derived atom-centered charges for each participating monomer.[11]

Thus, for obtaining the optimum configuration for the base pair XY (e.g. WCAT/GC pair), the critical points (CPs)[15] in X and Y are located and identified. The MESP-derived atom-centered charges for these species are also determined. The knowledge of MESP topography is used for positioning X and Y such that the hydrogen atoms in X are close to MESP minima in Y and vice versa. Either X or Y is held fixed and the other is allowed to translate and rotate with respect to the relevant degrees of freedom until a minimum is found for the interaction energy expressed by $E = {}^1/_2\{\sum V_{A,i}q_{B,i} + \sum V_{B,i}q_{A,i}\}$, V being the MESP of one species evaluated at the ith atomic site of the other species, where the potential-derived charge is q. Use of vdW radii of heavy atoms and appropriately scaled hydrogen radii prevents the collapse of the two species.

The procedure adopted for constructing the trimer geometries using the EPIC model is as follows: Let XY be a WC base pair (which is optimized with EPIC model/SCF level) that may bind to a base Z with either Ht or reverse Ht (Rt) hydrogen bonding between Y and Z. For applying the EPIC model to a trimer XYZ, XY is treated as a monomer for docking with Z. A block diagonal density matrix for XY is constructed with the density matrixes of X and Y. This is used for carrying out topographical analysis of MESP and determining MESP-derived charges for XY. The error introduced by such an approximation is checked by using the ab initio SCF-optimized geometry (planar constrained) for XY and the corresponding density matrix.

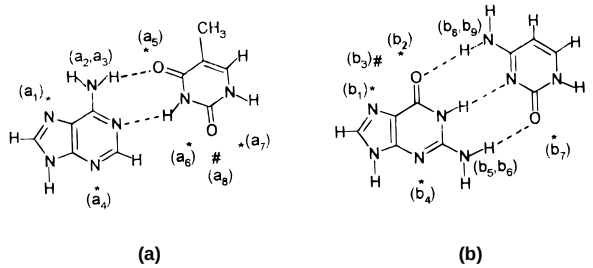

Figure 1. (a) AT and (b) GC Watson−Crick base pairs. * denotes (3,+3) MESP minima, and the # symbol denotes a (3,+1) saddle point in MESP. Bond lengths are not to scale (cf. Table 1).

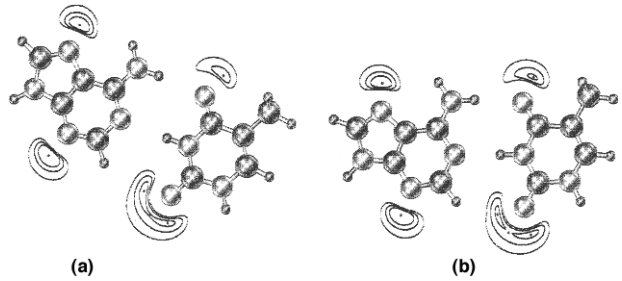

Figure 2. MESP contour plots in the ring plane for the AT base pair obtained with (a) the EPIC model and (b) ab initio optimization. The contour values are −0.05, −0.07, and −0.085 au successively toward the CPs marked by dots.

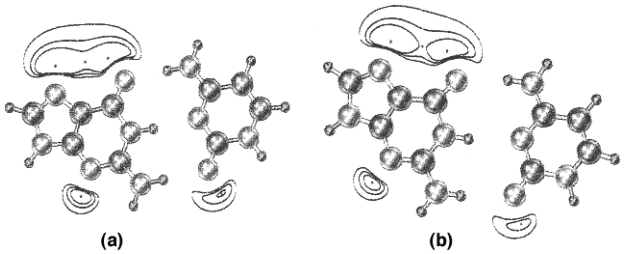

Figure 3. MESP contour plots in the ring plane for the GC base pair obtained with (a) the EPIC model and (b) ab initio optimization. The contour values are −0.05, −0.07, and −0.085 au successively toward the CPs marked by dots.

The geometries of A, T, G, and C are optimized with a constraint of planarity, at the HF/6-31G** level using GAMESS.[16] The planar optimized (6-31G**) structures of DNA base pairs AT and GC used in this work have been made available by Dr. J. Sponer. The MESP topographical features of molecules are stabilized[17a] at this level of basis set. Further, it has been shown that the incorporation of electron correlation does not alter the CP characteristics[17b] significantly at this basis. The UNIPROP[18] and GRID[19] programs are used respectively to carry out topographical analysis of MESP and obtain the MESP-derived atom-centered charges for the bases as well as their dimers. The optimization is carried out with a cyclic relaxation procedure, while simulated annealing technique is used for trapping the local minima. The graphical structures illustrating planar contour maps and the trimer orientations depicted in Figures 2, 3, and 4, respectively, have been obtained with the program UNIVIS.[20]

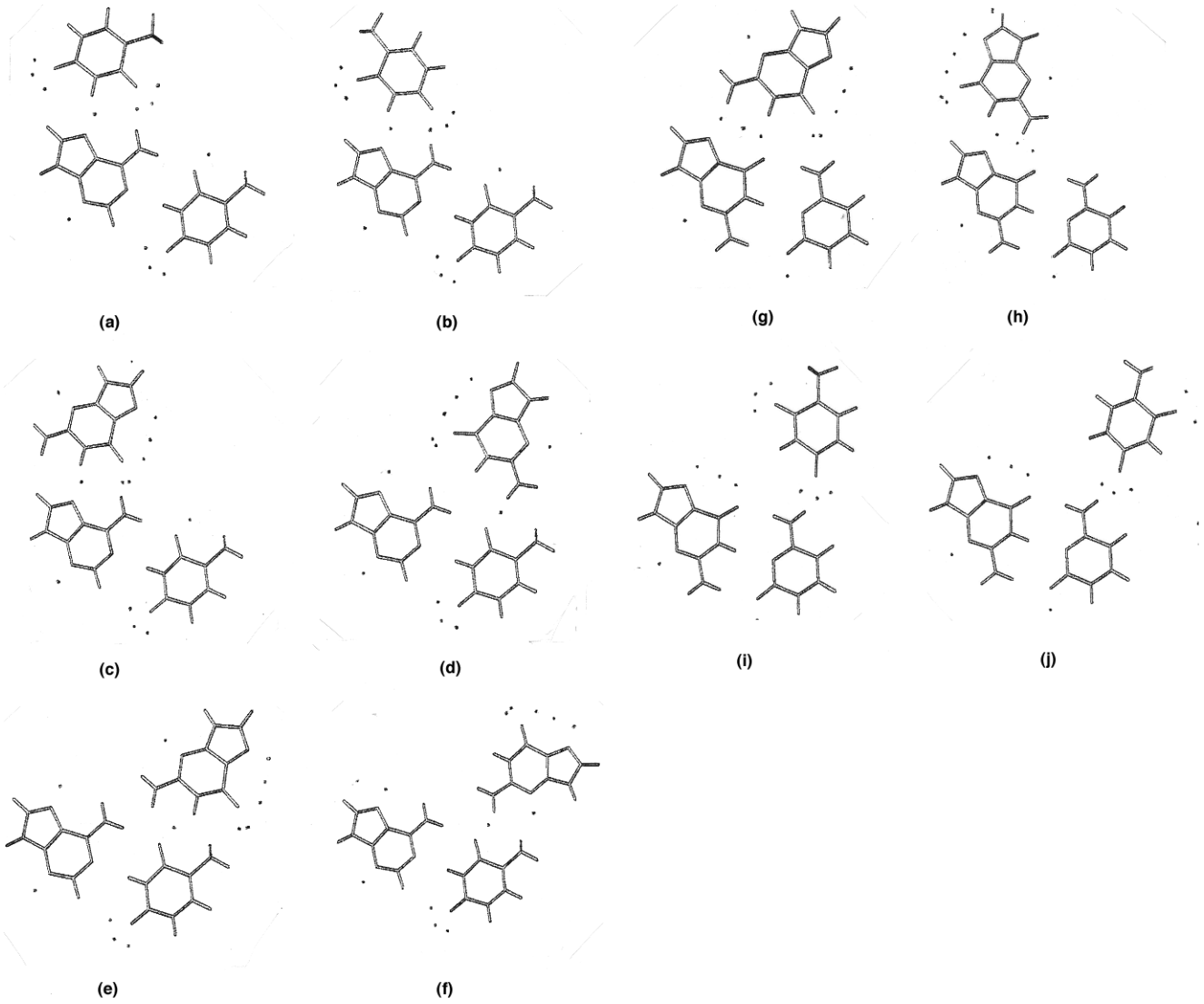

Figure 4. Geometries of various DNA trimers obtained with the EPIC model. The dots represent the critical points in either a WC dimer or the monomer. (a) TATHt, (b) TATRt, (c) TAG, (d) ATG, (e) ATGa, (f) ATGb, (g) CGGHt, (h) CGGRt, (i) TCGa, and (j) TCGb.

3. Results and Discussion

The geometrical orientations of the WC AT and GC base pairs are shown schematically in Figure 1a and b, respectively. The EPIC model interaction energies for these using the 6-31G** molecular wave function are −10.51 and −22.32 kcal mol^{-1}, respectively. Table 1 describes the MESP topographical features of these two complexes formed with the present model (block diagonal density matrixes used) and those optimized at the HF/6-31G** level. The CPs[15] in MESP representing oxygen and nitrogen lone-pairs that lie in the plane of aromatic rings are marked in Figure 1. The off-plane pair of minima associated with the amino nitrogen are indicated alongside the amino group. Most of these CPs are also seen (marked by dots) in Figures 2 and 3, wherein the contour maps of MESP are plotted in the ring plane for the two WC dimers. In Figures 2a and 3a, the MESP plots are obtained by using an additive density matrix for the model-predicted structure of AT(GC), while an actual dimer density matrix for the ab initio optimized structure is used for the plots in Figures 2b and 3b. The a and b parts of Figure 2 and also Figure 3 compare very well with respect to the negative MESP features, stressing the similarity in the respective charge concentration patterns. It may be predicted that such regions of negative MESP make these dimers susceptible to further interaction with other bases.

Examination of Table 1 reveals that, on pairing, the CPs that are not in the vicinity of the hydrogen-bonded region with the complementary base acquire more negative MESP values as compared to the monomer ones. For example, the CP marked (a_1) in Figure 1a for AT attains a slightly more negative MESP value of −0.103 au by modeling and −0.101 au after optimizing

TABLE 1: MESP Topographical Features in AT and GC Watson−Crick Base Pairs[a]

base pair	CP and its nature	MESP at CP in monomer	MESP at CP with EPIC model geometry	MESP at CP with SCF optimized geometry
A·T	(a_1) (3,+3)	−0.097	−0.103	−0.101
	(a_2,a_3) (3,+3)	−0.029	−0.043	−0.038
	(a_4) (3,+3)	−0.105	−0.110	−0.109
	(a_5) (3,+3)	−0.084	−0.079	−0.087
	(a_6) (3,+3)	−0.086	−0.092	−0.086
	(a_7) (3,+3)	−0.084	−0.089	−0.089
	(a_8) (3,+1)	−0.080	−0.086	−0.085
G·C	(b_1) (3,+3)	−0.125	−0.127	−0.134
	(b_2) (3,+3)	−0.120	−0.111	−0.125
	(b_3) (3,+1)	−0.085	−0.084	−0.094
	(b_4) (3,+3)	−0.084	−0.099	−0.105
	(b_5,b_6) (3,+3)	−0.005	−0.041	−0.040
	(b_7) (3,+3)	−0.111	−0.080	−0.087
	(b_8,b_9) (3,+3)	−0.015	−0.038	−0.024

[a] The subscripted letters associated with the CPs are marked in Figure 1. MESP CP values observed in purine and pyrimidine monomer bases are also given for comparison (all values in au). See text for details.

as compared to the adenine MESP value of −0.097 au. The decrease in the negative potential is substantial for the off-plane amino minima, marked (a_2,a_3) in Figure 1a and (b_8,b_9) in Figure 1b. For the two minima (b_5,b_6), however, this decrease is dramatic, changing the MESP value associated with the nitrogen lone pair from −0.005 au in guanine to −0.041 au in the GC pair (cf. Figure 1b). This will have significant repercussions for stacking interactions of two base pairs. The electrostatic interactions that are responsible for stacking of bases[21] will be much more favorable in base pairs, resulting in substantially different binding patterns for the latter as compared to the former ones. It may even be expected that there is a stronger electrostatic driving force for the formation of a duplex oligomer as compared to the single-stranded chain.

As far as trimers are concerned, binding of the third base with a dimer is expected to be stronger than that with an isolated base at a given position. The similarity of the values in the third and fourth columns of Table 1 also brings out the fact that the use of a block diagonal density matrix for a dimer does not alter the topographical characteristics of MESP significantly, especially in the region that is away from the interaction zone of the dimer. Hence, for predicting the binding energy with a third base using an electrostatic model, such an approximation in the dimer wave function is not expected to introduce large errors and, at the same time, offers large savings in computational time.

The interaction energy for a trimer involves two contributions: one associated with the dissociation into a monomer and a dimer, the other with the dissociation of the dimer into two monomers. The latter quantity, ΔE_{MD}, is either −10.51 (AT) or −22.32 (GC) kcal mol^{-1} with the EPIC model. The trimers studied here include one or more structures each of TAT, TAG, ATG, CGG, and TCG. A triplex containing Ht type hydrogen bonding between the second and the third bases is represented by XYZHt, while XYZRt stands for reverse Ht orientation of the third base. ATGa and ATGb represent trimers with an unusual orientation of guanine with respect to the WC AT base pair. The values of interaction energy involved in a reaction XYZ → XY + Z (ΔE_{DT}) as well as the total interaction energy associated with XYZ → X + Y + Z (ΔE_{MT}) are reported in Table 2. The ΔE_{DT} values are similar for base pair configurations obtained by the EPIC model and the optimized ones. The ΔE_{MT} values may be compared with the interaction energies obtained by performing single-point SCF calculations with the 6-31G** basis, at the model geometries (for a trimer XYZ, $\Delta E_{SCF} = E_{XYZ} - E_x - E_y - E_z$, the first term in this expression being the energy of the trimer and the rest of the terms stand for the monomer SCF energies). The ab initio interaction energies without incorporation of the basis set superposition error (BSSE) along with the corresponding gradient norms are included in Table 2. Figure 4 depicts the orientations of triplexes obtained with the EPIC model.

The trimer TAT has been qualitatively found[3] to be one of the most stable triplexes inside the Py−Pu−Py-rich oligomers, using various experimental techniques including UV melting[5a] and measuring[5b] the amounts of products cleaved by Fe-EDTA. The molecular modeling studies on d(TAT)$_{27}$ by Cheng and Pettitt[1] have demonstrated that various opposing factors cause TATHt (Figure 4a) to have a stability similar to that of TATRt (Figure 4b). Our gas-phase results are comparable to these findings, although the energetic preference is slightly more toward the Rt bonding (cf. Table 2). The trend again changes marginally on performing ab initio calculations for the interaction energies at the model-predicted geometries. The difference in the binding energies of these two possible structures is very small considering the accuracy involved in the method of calculations. Various factors, viz. level of basis set, BSSE, correlation effects, may also alter the energetic preference to some extent. The components of the interaction energy, ΔE_{DT} and ΔE_{MD}, differ by a very small amount in this case, implying that the experimental finding[5d] of having a single melting temperature for this triplex may have a molecular origin.

TABLE 2: Interaction Energies for the Formation of a Trimer from the Dimer (ΔE_{DT}) and those for the Formation of a Trimer from the Monomers (ΔE_{MT}) for Various DNA Triplexes Studied in This Work[a]

base triplex	ΔE_{DT}	ΔE_{MT}	ΔE_{SCF}	max. norm of gradient
TATHt	−9.34 (−10.44)	−19.85	−20.90	0.019
TATRt	−10.19 (−10.18)	−20.70	−20.48	0.020
TAG	−8.69 (−8.78)	−19.20	−19.13	0.024
ATG	−4.74 (−4.72)	−15.25	−14.69	0.041
ATGa	−4.68 (−4.42)	−15.19	−15.90	0.036
ATGb	−2.31 (−2.40)	−12.82	−14.52	0.035
CGGHt	−14.44 (−11.68)	−36.76	−41.85	0.040
CGGRt	−13.41 (−13.0)	−35.73	−38.47	0.034
TCGa	−6.25 (−6.58)	−27.57	−28.30	0.040
TCGb	−7.22 (−6.8)	−29.54	−30.58	0.031

[a] The single-point HF/6-31G** values (ΔE_{SCF}) at the EPIC model geometries and the corresponding gradient norms are displayed in the last two columns. The values in parentheses in the second column are obtained with optimized base pair geometries (all interaction energies in kcal mol^{-1}, the gradient norm in au). See text for details.

For the TAG (Figure 4c) and ATG (Figure 4d) trimers, most of the experimental[3] data predict the latter to be more stable than the former. The gas-phase electrostatic interactions in TAG for the configuration shown in Figure 4c are favorable for two hydrogen bonds imparting stability to the structure. Such unusual triads occurring in a Py−Pu−Py triplex were first proposed by Johnson and Morgan;[5f] however, recent studies[5a−e] indicated that such a structure is not easily accommodated in the triplex because of the unfavorable orientations of the glycoside bonds of the third strand.[3] The experimentally observed binding capability in ATG could be more due to steric as well as sequence dependent factors.[4e] For instance, it is possible that by adopting an unusual orientation, the third strand may impart an overall stability to the oligomeric structure. The stability is also influenced by the neighboring triplets so that a different choice of assay systems may lead to different inferences.[4e] In a 32-base ATG oligonucleotide, such an unusual pairing between second and third strands has been observed with NMR[4e] data, involving a single hydrogen bond between the thymine oxygen and the guanine amino proton. Out of the two such unusual patterns, ATGa and ATGb (shown in Figure 4e and f, respectively), the former was suggested by earlier workers.[5d] More recent studies by Wang et al.[4e] indicate that the NMR data are also consistent with the configuration ATGb, which should be more suitable from the point of view of the overall stability of the helical structure. Energy minimization studies[5e] of the triplet structures formed by heptameric strands with a mismatch Ht pairing at the central position have revealed that G in ATG is least unfavorable as compared to the other mismatch bases since, apart from the H-bond between its amino proton and T oxygen, there is yet another loose H-bond with T oxygen that lies above or below the ATG triplet. In fact, G is found to be appreciably tilted[5e] with respect to the average plane of this triplex. The EPIC model trimer interaction energies for both these forms are numerically quite small, with ATGa slightly more stable than ATGb, again demonstrating clearly the fact that the sequence specific interactions are also responsible for the triplex stability as pointed out by Wang et al.[4e] Also, the observation of a single melting temperature for ATG[4e] does not

seem to have a molecular origin, since the interaction energies for dimer and trimer formation (ΔE_{MD} and ΔE_{DT}) are quite different in this case.

The present model predicts CGG to be the most stable triplex. This finding is consistent with the earlier works.[1,5e] The major contribution to this stability comes from the interaction energy of the GC pair (-22.32 kcal mol^{-1}) involving three weak H-bonds. In both CGGHt and CGGRt, two hydrogen bonds are seen to exist (cf. Figure 4g,h) between the second and the third motif. In a 27-mer, molecular modeling[1] studies show CGGHt to be more stable than the reverse Ht bonding, in line with the present findings (cf. Table 2). The EPIC model also predicts Ht orientation to be somewhat more stable because of favorable interaction between amino protons in cytosine, the first monomer, and the CPs near oxygen in guanine, the third monomer. Such an interaction is missing in the Rt orientation. This structure has been ruled out, however, on the basis of very small solvation free energy for CGGHt. UV and NMR melting experiments[2] reveal a very high temperature for a single-stranded triple repeat sequence containing CGG. Similarly, half-dissociation temperatures for a 13-mer strand binding to a 31-mer duplex show that CGG is the next stable triplet after CGC.[5e] This observation is also consistent with the earlier reported molecular modeling studies.[5d] Thus, it seems that the favorable interactions between the three bases in CGG persist in the oligomeric form as well.

Two possible stable structures of the triplex TCG as obtained by the EPIC model are depicted in Figure 4i,j. A structure similar to the fomer has been observed[3,4d] to be stabilized by a single hydrogen bond involving the thymine oxygen and the cytosine amino proton in an unusual triad. Also, such a motif is easily accommodated within a Py−Pu−Py type triple-stranded structure.[4d] In an oligomeric sequence containing Pu−Pu−Py-rich strands, the TCG triplex unit has been observed[4c] in the central position, which occurs due to GC inversion. The geometrical disposition of the three units in this case[4c] (Figure 4i) too is similar to that stipulated by Radhakrishnan and Patel[4d] from their NMR data, although the perturbation caused by GC inversion has been found to influence two layers of triplexes on either side.[4c] In our study, we find that another orientation of the thymine unit as shown in Figure 4j also leads to a possible structure somewhat more stable than the one shown in Figure 4i, although this need not be necessarily suitable for imparting stability to the oligomeric triplex.

It is thus evident that the inherent binding capability in DNA trimers estimated with an electrostatic approach sometimes persists in the long chains, although it may be assisted or restricted by factors other than the molecular ones. The EPIC model gives quick predictions of stabilities (∼6 CPU hours of HP 715 workstation for modeling a triplet), and that these are quite reliable is proved by the values depicted in the last two columns in Table 2. The ab initio SCF interaction energies (ΔE_{SCF}) for trimer formation are quite congruous with the model values ΔE_{MT}, the largest deviation of ∼13% occurring for ATGb and CGGHt. The maximum value of the gradient norm for the model-predicted configurations is at the most 0.04 au, which is quite small, considering the approximate nature of the model (monomer units are rigid) and the time spent in obtaining these structures.

4. Conclusions

The gas-phase interaction energies of different DNA base trimers formed by the WC AT and GC base pairs are determined by the EPIC model. These dimers are treated as monomers in the model and are faithfully represented by the diagonally additive density matrixes of the respective monomers, as is evident from the comparison of MESP features in thus synthesized pairs and the ab initio optimized WC pairs. The locations of MESP minima in a base pair that are away from the weak binding region are not altered; however, the MESP values there become more negative as compared to those in the monomers. Consequently, the DNA base pairs are inherently more capable of forming further H-bonded structures. There is a substantial deepening of MESP values at the off-plane CPs accompanying amino nitrogens, the change being an order of magnitude in the case of guanine. This might result in dissimilar stacking features in bases and base pairs.

The trends in the stabilization energies obtained in the present work are according to those observed by the experimental and/or molecular dynamics studies in the case of TAT and CGG. The latter techniques tend to incorporate most of the in vivo conditions. Hence, it is presumed that the stabilities of these two triplexes originate at the molecular level. For the triplexes TAG and ATG, the gas-phase predictions are different than the observed ones, signifying the importance of the surrounding environment in these cases.

In general, the EPIC model interaction energies agree very well with the corresponding single-point SCF (6-31G**) values at the model orientations. This together with a small value of the maximum gradient norm (∼0.03 on average) illustrates that the model structures may not be too different from the fully optimized ab initio ones. The effects due to the type of basis set, BSSE, level of theory, correlation, etc., will surely influence the binding features to some extent. The qualitative behavior in terms of relative stabilities is not, however, expected to vary much. Whereas it takes only a few hours on a HP 715 workstation to scan the potential energy surface with the EPIC model, complete ab initio optimization even for base pairs requires a few hundred Cray CPU hours. The use of such an electrostatic modeling is, therefore, advantageous for knowing the recognition patterns as well as for comparing relative stabilities of different trimers in the gas phase. This approach also has a potential for investigating even larger molecular aggregates leading to self-assembled structures. Work is in progress in this direction in our laboratory.

Acknowledgment. The authors thankful to Dr. J. Sponer for providing optimized geometries of base pairs as well as for helpful discussions on the internet. Thanks are due to Council of Scientific and Industrial Research (CSIR), New Delhi, for providing financial support.

Note Added in Proof: The optimized structures of various base trimers have been recently reported: Sponer, J.; Burda, J. V.; Mejzlik, P.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dynam.* **1997**, *14*, 613. The interaction energies for TAT, TAG, and CGG in the present work agree quite well with the values for fully optimized structures reported in this work.

References and Notes

(1) Cheng, Y.; Pettitt, B. M. *J. Am. Chem. Soc.* **1992**, *114*, 4465.

(2) Zheng, M.; Huang, X.; Smith, G. K.; Yang, X.; Gao, X. *J. Mol. Biol.* **1996**, *264*, 323.

(3) Soyfer, V. N.; Potaman, V. N. *Triple-Helical Nucleic Acids*; Springer: Berlin, 1996.

(4) (a) Radhakrishan, I.; Santos, C.; Patel, D. I. *J. Mol. Biol.* **1993**, *234*, 188. (b) Rana, V. S.; Barawkar, D. A.; Ganesh, K. N. *J. Org. Chem.* **1996**, *61*, 3578. (c) Dittrich, K.; Gu, J.; Tinder, R.; Hogan, M.; Gao, X. *Biochemistry* **1994**, *33*, 4111. (d) Radhakrishan, I.; Patel, D. J. *Biochemistry* **1994**, *33*, 11405, and references therein. (e) Wang, E.; Malek, S.; Feigon, J. *Biochemistry* **1992**, *31*, 4838.

(5) (a) Miller, P. S.; Cushman, C. D. *Biochemistry* **1993**, *31*, 2999. (b) Best, G. C.; Dervan, P. B. *J. Am. Chem. Soc.* **1995**, *117*, 1187. (c) Chandler, S. P.; Fox, K. R. *FEBS Lett.* **1993**, *332*, 189. (d) Griffin, L. C.; Dervan, P.

B. *Science* **1989**, *245*, 967. (e) Mergny, J. L.; Sun, J. S.; Rougee, M.; Garestier, T. M.; Barcelo, F.; Chomilier, J.; Helene, C. *Biochemistry* **1991**, *30*, 9791. (f) Johnson, D.; Morgan, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 1637.

(6) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737.

(7) (a) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 1965. (b) Sponer, J.; Florian, J.; Hobza, P.; Leszczynski, J. *J. Biomol. Struct. Dynam.* **1996**, *13*, 827. (c) Colominas, C.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1996**, *118*, 6811. (d) Hobza, P.; Hubalek, F.; Kabelac, M.; Mejzlik, P.; Sponer, J.; Vondrasek, J. *Chem. Phys. Lett.* **1996**, *257*, 31, and references therein.

(8) (a) Allinger, N. L.; Kok, R. A.; Imam, M. R. *J. Comput. Chem.* **1988**, *9*, 591. (b) Leszczynski, J. *J. Phys. Chem.* **1992**, *96*, 1649. (c) Zahradnik, R.; Hobza, P. *Int. J. Quantum Chem.* **1986**, *29*, 663. (d) Sponer, J.; Hobza, P. *Int. J. Quantum Chem.* **1996**, *57*, 959.

(9) Pullman, B.; Claverie, P.; Caillest, J. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *57*, 1663.

(10) (a) Ghio, C.; Tomasi, J. *Theor. Chim. Acta (Berlin)* **1973**, *30*, 151. (b) Alagona, G.; Tani, A. *J. Chem. Phys.* **1981**, *74*, 3980.

(11) Gadre, S. R.; Bhadane, P.; Pundlik, S. S.; Pingale, S. S. In *Molecular Electrostatic Potentials: Theory and Applications*; Murray, J. S., Sen, K. D., Eds.; Elsevier: Amsterdam, 1996.

(12) Gadre, S. R.; Pundlik, S. S. *J. Phys. Chem.* **1997**, *101*, 3298.

(13) (a) Politzer, P.; Truhlar, D. G. *Chemical Applications of Atomic and Molecular Electrostatic Potentials*; Plenum: New York, 1981. (b) Tomasi, J.; Bonaccorsi, R.; Cammi, R. In *Theoretical Methods of Chemical Bonding Vol. 3*; Maksic, Z. B., Ed.; Springer: Berlin, 1990. The molecular electrostatic potential (MESP), V, at a point $\mathbf{r}$ is defined as

$$V(\mathbf{r}) = \sum_{A} \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}|} - \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|} d^3r'$$

where $\{Z_A\}$ are charges at nuceli situated at $\{R_A\}$ and $\rho(\mathbf{r})$ denotes the molecular charge density. The form of this expression suggests that $V(\mathbf{r})$ will vary between positive and negative limits and can provide useful information regarding electron-rich sites.

(14) (a) Buckingham, A. D.; Fowler, P. W. *J. Chem. Phys.* **1983**, *79*, 6426. (b) Dykstra, C. E. *Chem. Rev.* **1993**, *93*, 2339. (c) Alhambra, C.; Luque, F. J.; Orozco, M. *J. Phys. Chem.* **1995**, *99*, 3084.

(15) Gadre, S. R.; Kulkarni, S. A.; Shrivastava, I. H. *J. Chem. Phys.* **1992**, *96*, 5253. The topographical analysis of a scalar field like MESP, $V(\mathbf{r})$, involves locating and characterizing the points in space around a molecule, where the first space derivatives of this field vanish. Such points are termed critical points (CPs). The nature of such CPs is known by examining the eigenvalues of the corresponding Hessian matrix. Thus, a nondegenerate CP (rank of the Hessian equal to 3) could be any one of (3,+3) (minimum), (3,−3) (maximum), a saddle of the type (3,+1) or (3,−1). In the designation (R,s) for a CP, "R" represents the rank of the Hessian and "s" denotes the sum of signs of its eigenvalues.

(16) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. J.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *GAMESS. J. Comput. Chem.* **1993**, *14*, 1347.

(17) (a) Gadre, S. R.; Kulkarni, S. A.; Suresh, C. H.; Shrivastava, I. H. *Chem. Phys. Lett.* **1994**, *239*, 273. (b) Kulkarni, S. A. *Chem. Phys. Lett.* **1996**, *254*, 268.

(18) Shirsat, R. N.; Bapat, S. V.; Gadre, S. R. *UNIPROP*. *Chem. Phys. Lett.* **1992**, *200*, 373.

(19) Chipot, C. *GRID*: The FORTRAN program for fitting charges to molecular electrostatic potentials and fields; University de Nancy I: France, 1992.

(20) Limaye, A. C.; Inamdar, P. V.; Dattawadkar, S. M.; Gadre, S. R. *UNIVIS*: a visualization package, *J. Mol. Graph.* **1996**, 14.

(21) Gadre, S. R.; Pundlik, S. S. Unpublished work.