Indian Journal of Chemistry
Vol.39A, Jan-March 2000, pp. 132-138

Effect of water continuum on the interaction energy of DNA base- pairs: A Hartree-Fock self-consistent reaction field study

D Sivanesan, V Subramanian*[†], Balachandran Unni Nair & T Ramasami
Chemical Laboratory, Central Leather Research Institute, Adyar, Chennai, 600 020, India

Received 29 September 1999; accepted 29 November 1999

The stability of several base-pairs has been calculated in water medium using self-consistent reaction field theory (SCRF) in the framework of Hartree-Fock (HF) theory formalism. The nomenclature used by Sponer *et al.* (*J phys Chem*, 100(1996) 1965) have been used for all the base- pairs investigated in this work. The binding energy is found to increase for TAH, GT2, GC1, GA4, and GG3 base-pairs whereas, it decreases for GCWC, GCNEW, TARH, AC1, GT1, and TT1 base- pairs in solvent environment. The polarizing cavity around the base -pairs influences the primary electrostatic contribution arising from dipole-dipole interaction and charge distribution.

Introduction

It is well known that the three dimensional structure of DNA is influenced by the solvent environment[1-3]. Transition from one form of DNA to another is influenced by solvent, salt and polarity of media. The binding of small molecules by DNA is modulated by the solvent environment and presence of counter-ions[1,2]. Water molecules could provide dielectric screening between electrically charged entities and therefore influence electrostatic interaction in DNA. The solvent can control short-range hydrogen bonding interaction and long-range solvent polarization interaction. X-ray crystallography provides information about the possible roles of solvent wherein the solvent structure around DNA is ordered. However, it is a difficult task to estimate experimentally the energies of interaction of solvent and counter-ions with DNA.

Proper inclusion of solvent in theoretical calculation is a difficult task. Numerous researchers have applied classical molecular dynamics (MD) or Monte Carlo (MC) simulation methods to predict the role of solvent in the stabilization of three dimensional structure of DNA[4-17]. One can use continuum models based on Poisson-Boltzmann equation and quantum chemical studies employing various variants of Self Consistent Reaction Field (SCRF) approach at the semi-empirical or *ab initio* level. Though all these approaches have merits and demerits, at least qualitative inclusion of solvent effect is necessary since results derived from gas phase data can seldom be complete.

With the advent of fast computers and computational methodologies, it is now possible to perform high-level quantum chemical calculations on the isolated DNA bases and base-pairs. Status of quantum chemical *ab initio* studies on hydrogen- bonding and stacking of DNA bases has been reviewed by Sponer *et al*[18]. Several high level *ab initio* calculations have been performed to unravel the structure and properties of isolated DNA bases and hydrogen- bonded base-pairs[19-27]. The energetics of base-pairing of DNA bases have been calculated using electrostatic potential for intermolecular complexation model[28,29]. These results are obtained from gas phase calculations. Energy needs for intra-molecular proton transfer in mono and dihydrated tautomers of guanine have been calculated using *ab initio* post Hartree-Fock theory[30a]. Recently, using density functional method, water-assisted intramolecular proton transfer in the tautomers of adenine has been carried out by Gu and Leszczynski[30b]. Energetics of reverse Watson-Crick base-pairing have been calculated for isocytosine-cytosine and guanine-cytosine base-pairs in the gas phase and in a water cluster using ab initio quantum chemical methods[31]. These calculations have confirmed that water influences the structure, proton transfer rate and the stability of DNA bases in solution.

Although there are several methods to treat solute-solvent interactions, the Onsager method provides the simplest approach based on uniform continuum dielec-

---

[†] email: clrimd@vsnl.com,subuchem@hotmail.com

tric sphere around the solute[32,33]. The permanent dipole moment of solute will induce a dipole in the medium, which inturn will interact with the molecular dipole leading to its stabilization. In the SCRF-MO formalism, the solute-solvent interaction is treated as perturbation of the Hamiltonian of the isolated molecules[34,35]. The free energy of solvation, estimated using the SCRF method, accounts for electrostatic interaction between solute and solvent. Recently, we have assessed the role of solvent in the stacking of cytosine dimers using Onsager reaction field model in the framework of Hartree-Fock formalism[36]. Though the continuum solvation model is not adequate to explain exact explicit solvation of base-pairs, it is a useful method to obtain information about influence of solvent on the energetics of base-pairing and base-stacking.

Two distinct mechanisms are responsible for hydration of DNA. Depending on the extent of hydration, few water molecules interact explicitly with base-pairs. The continuum formed by the water molecules could also influence the structure and stability of base-pairs and stacked bases. In this study, an attempt has been made to understand the question as to how the solvent continuum influences the stability of base-pairs. The relative stability of base-pairs in a solvent environment has been assessed and a comparison has been made with those obtained from gas phase calculations. Previous *ab initio* study on the structure and energies of base-pair reveals that, the primary stabilization of the H-bonded DNA base-pairs originate from the electrostatic interaction, which can be treated effectively within the HF approach[19]. Hence, in the present study, an attempt has been made to emphasize the importance of solvent effect and the binding energies of several base-pairs has been carried out in the framework of HF theory by applying SCRF approach. The SCRF method lacks specific interaction of water with solute molecules and a level of truncation of electrostatics is very important. Hence, calculation has also been performed on Guanine-Cytosine Watson-Crick (GCWC) base-pair by including various multipoles in the SCRF calculation.

Methodology

Sponer *et al.* have carried out *ab initio* calculations for all possible combinations of neutral major tautomers of DNA bases such as adenine (A), cytosine (C), guanine (G) and thymine (T)[19]. The HF/6-31G**- optimized geometries for various planar base-pairs acquired from Prof. J. Sponer have been used without further optimization (single point calculations) to estimate the effect of water continuum on the binding energy of base-pairs. The base-pairs studied in this investigation are shown in Fig.1. The nomenclature and the hydrogen-bonding pattern adopted by Sponer *et al.* have been employed in this study[19]. HF/6-31G** and SCRF-HF/6-31G** single point calculations were performed on DNA base-pairs with the G94W suite of programs[37]. The interaction energy is calculated as V_{int} = E(base-pair) - E(base1)- E(base 2). Here, E(base-pair), E(base1) and E(base 2) are the HF total energies of base-pair and the individual DNA bases in both gas and solvent phases. The binding energy is equal to $-V_{int}$. The interaction energy was corrected for basis set super position error (BSSE) using Boys-Bernardi method[38]. The cavity sizes of base-pairs have been calculated using VOLUME keyword. The dielectric constant of water used in the calculation is 78.50. UNIVIS package was used to visualize the structures of the base-pairs[39]. MESP derived charges has been computed using CHELPG method.

Results and Discussion

The interaction energy obtained from previous study has been corrected for deformation energy (deformation energy of bases upon complexation was calculated as the difference between energies of planar optimized bases and energies of bases having the same geometry as within the complex) and zero point energy with correlation at MP2 level[19]. It is evident from the study that the electrostatic energy contributes significantly to the stabilization of the H-bonded base-pairs with large dipole moments. However, electron correlation is also important. For GCWC base-pair, correlation energy and deformation energy contributions are about 5% and 8%, respectively. Since electron correlation energy contribution is very small for the strong base-pairs, in this study changes in the electrostatic contributions upon solvation of base-pairs have been assessed.

Total energies and dipole moments of the various base-pairs have been calculated and presented along with polarization energy in Table 1. In general, the dipole moment of the base-pairs increases with polarity of the medium. As seen from Table1, the dipole moment of all the base-pairs differs significantly from each other. Therefore, stabilization of DNA base-pair in water solution is expected to vary with solvent. The Watson-Crick, reverse Watson-Crick, Hoogsteen and reverse Hoogsteen TA and GC pairs were designated using abbreviations WC, RWC, H and RH, respectively. The remaining pairs were numbered in order of decreasing base-pair stability obtained by lower level *ab initio* calculations. The

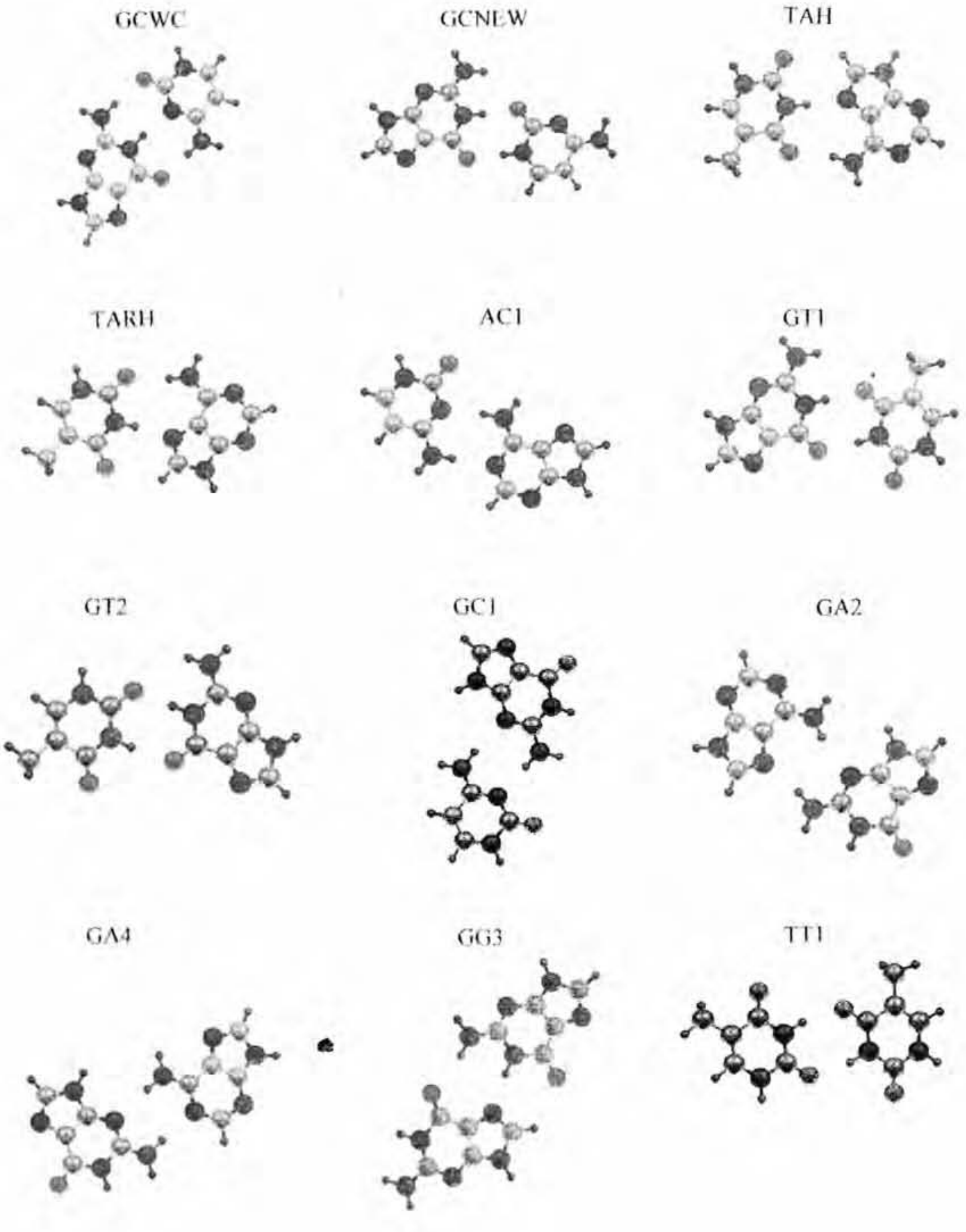

Fig.1 - The structures of base-pairs (the coordinates from ref.19)

Table 1 - Total energy (in Hartrees), polarization energies and dipole moment (Debye) at the HF/6-31G** optimized geometries

Base-Pair	Gas Phase		Solvent Phase		
	Total Energy	Dipole Moment	Total Energy	Polarization Energy	Dipole Moment
GCWC	-932.08458824	6.5	-932.089712061	-0.0066	8.4
GCNEW	-932.07990491	3.1	-932.080989937	-0.0014	4.1
TAH	-916.08012199	6.4	-916.084523709	-0.0054	7.9
TARH	-916.08006407	5.9	-916.083883022	-0.0047	7.4
AC1	-857.18737724	4.8	-857.189884049	-0.0031	5.9
GT1	-990.95997435	7.7	-990.966260288	-0.0077	9.5
GT2	-990.95944559	8.6	-990.966402955	-0.0084	10.4
GC1	-932.06431472	12.7	-932.080297148	-0.0193	15.3
GA2	-1003.9613022	7.3	-1003.96626641	-0.0062	9.1
GA4	-1003.9634698	9.2	-1003.97296547	-0.0123	11.9
GG3	-1078.8529524	10.5	-1078.86328466	-0.013	13.3
TT1	-903.06470981	1.3	-903.064885010	-0.0002	1.6

calculated interaction energies for GCWC, GCNEW (here, the N1(pyrimidine) and N9(purine) interactions are ignored), TAH, TARH, AC1, TT1, GG3, GA2, GA4, GC1, GT1 and GT2 in gas and solvent environment have been presented in Table 2. Our results show that hydrogen-bonded complexes such as GCWC, GCNEW, TARH, AC1, GT1 and TT1 base-pairs are preferred in gas phase.

Table 2 - Binding energy (kcal/mol) of base-pairs obtained at the HF/6-31G** optimized geometries

Base-Pair	Binding Energy	
	Gas Phase	Solvent Phase
GCWC	25.5	21.2
GCNEW	22.7	15.9
TAH	11.3	11.7
TARH	10.9	8.7
AC1	11.9	7.9
GT1	14.1	13.6
GT2	13.7	14.1
GC1	12.7	16.8
GA2	7.5	7.8
GA4	8.8	11.5
GG3	16.8	17.8
TT1	9.1	6.6

This may be rationalized by the fact that, in the case of hydrogen bonded pairs, the atoms that are involved in creating hydrogen bonds with other solute molecules are prevented from forming H-bonding with water molecules and hence base-pair in water are less stable when compared to gas phase calculation. Since SCRF formalism lacks solute-solvent structure, the above argument is valid to some extent. There are a number of factors contributing to the stability of base-pairs in gas phase and in solvent medium. In addition to the number of hydrogen bonds between base-pairs, the mutual orientation of base-pairs and their dipole moments are the other important factors contributing to the stability of base-pairs. It is evident from this calculation and also from other reports that the polarity of the base-pairs imparts significant influence on the strength of the base-pairs. In order to estimate the importance of dipole moments in the stabilization of base-pairs, the individual dipole moments of base-pairs have been calculated using HF/6-31G** in both gas phase and in water solution. The calculated dipole moments of G, C, A and T in gas phase are 7.25, 7.27, 2.52 and 4.58 Debye, respectively. Due to change in the polarity of the medium, the dipole moment of the base-pairs increases in water medium. The values for G, C, A and T are 8.98, 9.15, 3.25 and 5.32 Debye respectively. Hence, the stability of the base-pairs in water solution is determined by the changes in the dipole moment values. It can be observed from Table 2 that the presence of water continuum increases the binding en-

Table 3 - Calculated charges of GCWC base-pair in both gas phase and water medium

Atom	Gas Phase		Solvent Phase	
	MPA	NPA	MPA	NPA
C2	0.9952	0.9900	0.9977	0.9924
N1	-0.7632	-0.7177	-0.7617	-0.7145
C6	0.2435	0.1533	0.2496	0.1657
C5	-0.3132	-0.4516	-0.3190	-0.4595
C4	0.7052	0.6207	0.7082	0.6234
N3	-0.8260	-0.7776	-0.8396	-0.7897
O2	-0.6775	-0.7822	-0.6808	-0.7869
N4	-0.7764	-0.8645	-0.7801	-0.8704
H1	0.3382	0.4546	0.3483	0.4615
H6	0.1966	0.2406	0.2184	0.2523
H5	0.1746	0.2505	0.1819	0.2547
H41	0.3146	0.4198	0.3156	0.4209
H42	0.4126	0.4819	0.4045	0.4744
O6	-0.6807	-0.7676	-0.6727	-0.7578
C2	0.8476	0.8350	0.8506	0.8344
N1	-0.8787	-0.7532	-0.8755	-0.7523
C4	0.9476	0.8051	0.9496	0.8098
N3	-0.7338	-0.7127	-0.7414	-0.7225
C5	0.6546	0.4653	0.6551	0.4675
C6	-0.0368	-0.1355	-0.0416	-0.1377
N2	-0.5225	-0.4916	-0.5207	-0.4871
C8	0.3311	0.2594	0.3234	0.2438
N7	-0.7465	-0.6621	-0.7489	-0.6652
N9	-0.8206	-0.9052	-0.8119	-0.8985
H1	0.4139	0.4824	0.4204	0.4898
H21	0.1649	0.2153	0.1403	0.2026
H22	0.3219	0.4501	0.3153	0.4439
H8	0.3222	0.4326	0.3220	0.4325
H9	0.3893	0.4649	0.3974	0.4723

Table 4 - Total energy of GCWC base pair using HF/6-31g* basis set

Approximations	Total energy
Gas Phase	-932.05073
Scrf=Dipole	-932.05563
Scrf=Quadrupole	-932.08119
Scrf=Octapole	-932.06594

ergy of TAH, GT2, GC1, GA4 and GG3. The difference between binding energy obtained from gas phase and solvent continuum calculation for TAH, GA2 and GT2 is small. These base-pairs involve combination of less polar base-pairs. It is an important point to consider that for strong polar bases the dipole-dipole interaction is very important. In fact for the strong polar systems, correlation energy is shown to be repulsive[19]. For other base-pairs the secondary interactions, arising from the atomic characteristics such as point charges and atomic dipole moments and correlation energy contribution would lead to stabilisation of base-pairs.

From the gas phase calculation, it is evident that GCWC base-pair is the most stable base-pair followed by other base-pairs[19]. The stability of GCWC pair is related to large dipole moments of G and C in contrast to the dipole moments of A and T. Similar observation is evident in the solvent environment. Solvated GC pair is the most stable base-pair with binding energy of 21.2 kcal/mol. The principle contribution to binding energy of base-pair arises from the attractive dipole-dipole interaction, which is weakened by the presence of water continuum. The charges calculated from Mulliken population analysis (MPA) and natural population analysis (NPA) schemes for GCWC in both gas phase and water media are shown in Table 3. It can be observed that the solvent does not have uniform influence on the charge distribution. The polarization of charges after solvation is small and depends on the charges in gas phase. The changes in the charge distribution of GCWC pair upon solvation are reflected by 29% increase in the dipole moment of the GCWC base-pair. Similar variation of charge distribution pattern has also been observed in other base-pairs considered in this study. It is observed from MESP derived charges for all atoms in GCWC pair that polarization of charges is more pronounced in solvent when compared to other methods.

Out of the four possible combinations of T and A predicted in the previous investigation, we have calculated the binding energies of TAH and TARH in both gas and solution phase. These base-pairs have similar interaction energies in gas phase with one weak O-H(C) hydrogen bond. Although their dipole moments are nearly same in both gas phase and water media, the binding energies obtained from the present SCRF calculation are different. There is no striking influence of water continuum on the interaction energy of TAH (only a small increase of 0.4 kcal/mol) whereas binding energy of TARH base-pair decreases by 2.2kcal/mol.

It is observed from the previous report that AC1 and GC1 pairs contain N...(H)N bonds which are longer than 3.0Å(ref.19). The GC1 pair exhibits significantly longer N3...H(N4) bonds, which are longer than 3.0A. It is evident from Table 1 that GC1 pair has large dipole moment when compared to all other possible combinations of DNA bases in both gas and solution phase. Due to variation in the dipole moment and charge distribution, these base-pairs undergo different levels of stabilization upon solvation. It can be noted that the binding energy of AC1 decreases whereas, that for GC1 increases in the presence of solvent. In the gas phase GT1 and GT2 pairs have the same geometrical parameters and binding energy[19]. The reversion of thymine does not influence the binding energy of these base-pairs[19]. These pairs do not undergo any drastic changes in binding energy upon solvation by the continuum formed by the water. The TT1 pair has less dipole moment compared to all possible combinations of A, C, G and T. The binding energy of TT1 base-pair decreases by 2.5 kcal/mol with reference to the gas phase binding energy.

The previous investigations have shown that the GA2, GG4 and AA3 base-pairs are weakly bonded with a significant part of the stabilization arising from electron correlation contribution[19]. In weak complexes (formed by less polar DNA bases) electron correlation contributes about 30-40% to the binding energy. Since the effect of electron correlation is not considered in the present calculation, the stabilization energy obtained for GA4 complex may not be a realistic estimate when compared to other base-pairs for which the electrostatic contribution is the primary factor contributing to their stability.

The limitation of SCRF calculation is mainly due to the selection of cavity size and level of truncation of electrostatics. This model is appropriate for relatively compact molecules, which can be treated using dipole in a spherical cavity. Since the base-pairs are not simple and compact molecules to be treated by dipole in a spherical cavity, we have calculated the total energy of GCWC base-pair using multipoles in a spherical cavity using HF/6-31G* method. The calculated total energies are presented in Table 4. It is evident from the results shown in the table that the relative energy of the GCWC pair increases while going from dipole in a spherical cavity approximation to multipoles in the spherical cavity. It is noted from the population analysis that due to the presence of multipoles in a spherical cavity the polarization of GCWC pair enhances when compared to that of dipole approximation. These results suggest that it is possible to account for electrostatic interaction between the two bases in water environment using the multipole in a spherical cavity approximation. Previous studies have shown that it is not possible to compare the results derived from explicit interaction of water with base-pairs[30,31]. This is due to the fact that the solvent continuum model lacks structure and hence explicit interaction of water molecule with base-pair can not be realized. This model helps to understand the effect of polar medium on the relative stability of various base-pairs.

Conclusion

The present study focuses on the important issue related to the stability of various base-pairs in water medium. The binding energy of base-pairs such as GCWC, GCNEW, TARH, AC1, GT1, and TT1 decreases in the water medium whereas for pairs viz. TAH, GT2, GC1, GA4 and GG3, it increases. It is observed from the work of Sponer *et al.* that for GCWC, GT1, GT2, GG3 and TT1, the electron correlation energy contribution ranges from 5-14% (ref.19) and hence conclusion derived from the HF-SCRF calculation is of reliable quality. For other base-pairs electron correlation ranges from 20-36% and hence the result needs to be interpreted carefully. The interaction energy obtained from this work has limitation due to the absence of electron correlation contribution and changes in the geometry of base-pairs upon solvation. The inclusion of deformation energy and electron correlation would further decrease the binding energy of base-pairs in the solvent environment when compared to gas phase results.

These calculations reinforce the earlier arguments that water plays an important role in the stabilization of DNA base-pairs. In the SCRF method, the stabilization energy is related to the square of permanent gas phase dipole moment and polarization energy of the base-pairs in the presence of solvent. It is also important to consider the fact that the size and orientation of the dipole moments in both gas phase and water medium could also modulate the binding energy. In addition to the primary dipole-dipole interaction, the secondary interactions arising from the changes in the charge distribution pattern and multipoles are also important in the stabilization of base-pairs in water continuum. Due to redistribution of charges in the solvent media, the charges on the atoms involved in forming hydrogen bond (between $H^{\delta+}...X^{\delta-}$) are perturbed thereby influencing the secondary interaction responsible for the stabilization of base-pairs. Since continuum model lacks solute-solvent local structure, it is not possible to quantify the exact nature of interaction between base-base-water. The estimation of interaction energy of few water molecules with selected base-pairs is in progress using *ab initio* methods.

Acknowledgement

This study is supported by CSIR research grant No. 80(00022)/96-EMR-II. One of the authors (VS) wishes to thank Professor Jiri Sponer, J. Heyrovsky Institute of Physical Chemistry, Czech Republic, for his valuable suggestions and providing geometries of various bases and Prof. S. R. Gadre of Pune University for his interest

in this work. The UNIVIS package developed by Prof. S. R. Gadre and co-workers was used to visualize geometries of various base pairs used in this investigation.

References

1 Saenger W, *Principles of nucleic acid structure* (Springer-Verlag , New York), 1984.

2 Westhof W & Beveridge D L in *Hydration of nucleic acids*, edited by W Westhof & D L Beveridge (Cambridge University Press, Cambridge), Vol.5, 1989.

3) Lee R S & Bone S, *Faraday Discussion*, 103(1996)59.

4 (a) Jayaram B & Beveridge D L, *Annu Rev Biophys Biomol Struct*, 25 (1996) 367. (b) Ravishankar G, Auffinger P, Langley D R, Young M A & Beveridge D L in *Reviews in computational chemistry*, edited by K B Lipkowitz and D B Boyd (VCH Publishers, New York), 1997.

5 Auffinger P & Westhof E, *Curr Opin struct Biol*, 8(1988)227.

6 (a) Dang LX, Pearlman D A & Kollman P A, *Proc Natl Acad Sci USA*, 87(1990)4630.

(b) Dang L X & Kollman P A, *J Am chemSoc*, 112(1990)503.

(c) Dang L X & Kollman P A, *J Am chem Soc*, 112(1990)5716.

7 Langlef J, Claverie P, Caron F & Boeuve J C, *Int J quant Chem*, 19 (1981)229.

8 Pohorille A, Burt S K & MacElroy R D, *J Am chem Soc*, 106(1984)402.

9 Pohorille A, Pratti L R, Burt S K, & MacElroy R D, *J Biomol Struct Dyn*,1(1984)1257.

10 Egan J T, Nir S, Rein R & Macelroy R, *Int J quant Chem; Quant Biol Symp*, 5(1978)433.

11 Kollman P A, *Chem Rev*, 93(1993)2395.

12 Kollman P A & Merz K M, *Acc chem Res*, 23(1990)246.

13 Cieplak P & Kollman P A, *J Am chem Soc*, 110(1988)3734.

14 Freier S M, Alkema D, Sinclair A, Neilson T & Turner DH, *Biochemistry*, 24(1985)4533.

15 Young M A, Jayaram B & Beveridge, D L , *J Am chem Soc*, 119(1997) 59.

16. Mazur A K, *J Am chem Soc*, 120 (1998)10928.

17 Spackova N, Berger I, Egli M & Sponer J, *J Am chem Soc*, 120(1998)6147.

18. Sponer J, Leszczynski J & Hobza P, *J biomol Struct Dyn*, 14(1996)117.

19. Sponer J, Leszczynski J & Hobza P, *J phys Chem*, 100(1996)1965.

20. Sponer J, Leszczynski J & Hobza P, *J phys Chem*, 100(1996)5590.

21. Aida M, *J theor Biol*, 130(1988)327.

22. Alhambra C, Luque F J, Gago F & Orozco M, *J phys Chem B*, 101(1997)3846.

23. Sponer J & Hobza P, *Int J quant Chem*, 57(1996)958.

24. Florian J & Leszczynski J, *J Am chem Soc*, 118(1996)3010.

25. Aida M, Kaneko M & Dupuis M, *Int J quant Chem*, 57(1996)949.

26. Sugiyama H & Saito I, *J Am chem Soc*, 100(1996)7063.

27. Sponer J, Leszczynski J & Hobza P, *J phys Chem, A* 101(1997)9489.

28. Gadre S R & Pundlik S S, *J phys Chem, B* 101(1997)3298.

29. Pundlik S S & Gadre S R, *J phys Chem, B* 101(1997)9657.

30. (a) Gorb L & Leszczynski J, *J Am chem Soc*, 126(1998)5024

(b) Gu J & Leszczynski J, *J phys Chem*, 103(1999)2744.

31. Zhanpeisov N V, Sponer J & Leszczynski J, *J phys Chem, A* 102(1998)10374.

32. Onsagar L, *J Am chem Soc*, 58(1936)1486.

33. *Exploring chemistry with electronic structure methods* by J B Foresman & Aeleen Frisch, 2nd Edition, (Gaussian Inc, Pittsburgh).

34. Wong M W, Frisch M J & Wiberg K B, *J Am chem Soc*, 113(1990)4776.

35. Tapia O & Goscinski O, *Mol Phys*, 29(1975)1653.

36. Subramanian V, Sivanesan D & Ramasami T, *Chem Phys Lett*, 290(1998)189.