Three-Center Hydrogen Bonds in DNA: Molecular Dynamics of Poly(dA)·Poly(dT)

V. Fritsch and E. Westhof*

Contribution from the Laboratoire de Cristallographie Biologique, Institut de Biologie Moléculaire et Cellulaire du C.N.R.S., 15 rue R. Descartes, 67084 Strasbourg Cedex, France.
Received May 3, 1991

Abstract: The program AMBER 3.0 has been used to generate molecular trajectories of several models of dA·dT oligomers. The simulations were done without explicit solvent molecules but with two different dielectric functions ($\epsilon(r) = 4r$ and a sigmoidal distance-dependent dielectric function $\epsilon(r) = \epsilon_{cal}$). Best comparisons between experimental data based on Raman, NMR, and X-ray studies and calculated results are obtained with $\epsilon(r) = \epsilon_{cal}$. A comparative survey of the behaviors of the Watson–Crick and three-center hydrogen bonds was made to evaluate the importance of such hydrogen-bond systems in the stability of dA·dT sequences. Depending on the dielectric function, the lifetimes of the Watson–Crick H-bonds are 10 to 50 times longer than those of the three-center H-bonds in the major groove of dA·dT. The activation energies relevant for the three-center H-bonds are of the same order of magnitude as those underlying the pseudorotational movements of the puckered deoxyribose sugars (<1 kcal mol^{-1}). Despite mean lifetimes and activation energies for the three-center hydrogen bonds of the same order of magnitude as those for the hydrogen bonds in simulations of liquid water, the autocorrelation functions tend, in 0.5 ps, to an average value of 0.5, indicating a high probability of occurrence. Thus, three-center hydrogen bonds appear more as geometrical consequences of the anomalous structures adopted by dA·dT homopolymers (high propeller twist and low roll angles) than as structurally governing factors.

Introduction

The homopolymer poly(dA)·poly(dT) has many unusual properties that distinguish it from other B–DNA polynucleotides. In solution, the helical repeat was determined to be 10.0 base pairs per turn, while it was found at 10.5 base pairs for random DNA.[1] In fibers, the homopolymer is not affected by environmental changes like humidity, cations, and salt concentration, which induce the B- to A-form transition for other sequences.[2] In addition, poly(dA)·poly(dT) cannot be reassociated into nucleosomes while poly[d(A–T)]·poly[d(A–T)] can.[3] Stretches of adenines on one strand have also been associated with the propensity of some DNA sequences to bend.[4] A possible explanation for these anomalous properties was proposed recently on the basis of two DNA dodecamer crystal structures with a large stretch of adenine residues on one strand and thymine residues on the other.[5] In these studies, it was argued that the unusual conformation is due to the high propeller twist at each A–T base pair, resulting in the formation of cross chain three-center hydrogen bonds along the major groove (i.e., each adenine N6 amino group, beside being involved in the Watson–Crick hydrogen bond with the acceptor thymine O4 on the opposite strand, interacts with the acceptor thymine O4 of the 5′ thymine base of the oligo(dT) strand). Such bonds are possible at every A–A step. In contrast, Yoon et al.[6] remarked that such hydrogen bonds cannot be formed with an alternating A–T sequence and that a smaller propeller twist is observed for such a sequence.

In this work, molecular mechanics minimizations and molecular dynamics simulations were performed with the AMBER 3.0 force field to study the formation and behavior of three-center hydrogen bonds in different homopolymer models. The mean lifetimes of such hydrogen bonds were also evaluated in comparison with the Watson–Crick ones. To analyze the stability of the various hydrogen bonds, simulations were performed at a series of temperatures defined by the mean kinetic energy of the system between 50 K and 300 K.

Methods

The molecular mechanics computations and the molecular dynamics simulations were carried out using the program AMBER 3.0[7] where the potential energy function has the following form:

$$E = \sum_{\text{bonds}} k_d(d - d_0)^2 + \sum_{\text{angles}} k_\theta(\theta - \theta_0)^2 + \sum_{\text{dihedrals}} \frac{V_n}{2}(1 + \cos(n\phi - \gamma)) + \sum_{\text{nonbonded}} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + 332\frac{q_i q_j}{\epsilon r_{ij}} \right) + \sum_{\text{H bonds}} \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right)$$

The all-atom force-field parameters proposed by Weiner et al.[7] were used in this potential function. Because of the presence of charges and highly polar groups in nucleic acids, electrostatic interactions take an important part in the force field. In this potential energy function, the electrostatic energy is given by Coulomb law, where q_i and q_j are the partial charges and ϵ the relative dielectric permittivity. The choice of ϵ is not easy and different solutions were proposed. One of these consists of the use of a distance-dependent dielectric function $\epsilon(r)$, which will mimic the solvent local structure and dielectric saturation effects on the electrostatic energy, at least qualitatively. In this work, two dielectric functions were used: the function $\epsilon(r) = 4r$ proposed by the program AMBER and the sigmoidal function suggested by Lavery et al.,[8] called ϵ_{cal}, that we added to AMBER.[9] Clearly, explicit water molecules with ions could be introduced in an all-encompassing simulation at the cost of a much longer set-up and computation time with possible coherence and convergence problems. Although ultimately such calculations have to be performed, the development of a fast and easily mastered computational approach for testing and screening nucleic acid models is still necessary and useful. It should be noted that, in this first approach, a Debye screening function is not introduced and the full charges of the AMBER dictionary are used.

The initial structures were first refined using a conjugate gradient energy minimizer, until the root-mean-square energy gradients were less

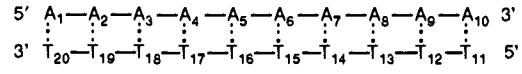
(1) Wang, J. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 200–203. Peck, L. J.; Wang, J. C. *Nature* **1981**, *292*, 375–378. Rhodes, D.; Klug, A. *Nature* **1981**, *292*, 378–380.
(2) Bram, S. *Nature New Biol.* **1971**, *232*, 174–176. Bram, S.; Tougard, P. *Nature New Biol.* **1972**, *239*, 128–131. Pilet, J.; Brahms, J. *Nature New Biol.* **1972**, *236*, 99–100.
(3) Rhodes, D. *Nucleic Acids Res.* **1979**, *6*, 1805–1816. Simpson, R. T.; Kuntzler, P. *Nucleic Acids Res.* **1979**, *6*, 1387–1415.
(4) Travers, A. A. *Annu. Rev. Biochem.* **1989**, *58*, 427–452. Hagerman, P. J. *Annu. Rev. Biochem.* **1990**, *59*, 755–781. Wu, H.-M.; Crothers, D. M. *Nature* **1984**, *308*, 509–513.
(5) Coll, M.; Frederick, C. A.; Wang, A. H. J.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8385–8389. Nelson, H. C. M.; Finch, J. T.; Luisi, B. F.; Klug, A. *Nature* **1987**, *330*, 221–226.
(6) Yoon, C.; Prive, G.; Goodsell, D.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 6332–6336.
(7) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230–252.
(8) Lavery, R.; Sklenar, H.; Zakrzewska, K.; Pullman, B. *J. Biomol. Struct. Dyn.* **1986**, *3*, 989–1014.
(9) Fritsch, V.; Westhof, E. *J. Comput. Chem.* **1991**, *12*, 147–166.

0002-7863/91/1513-8271$02.50/0

than 0.05 kcal mol^{-1} Å^{-1}. In order to analyze the dynamical properties of the molecule, molecular dynamics (MD) simulations were performed during 50 ps on the minimized structures obtained for each model. Although most simulations were done at 300 K, a set of simulations was calculated for a series of temperature (between 50 K and 300 K). The initial velocities were taken from a Maxwell-Boltzmann distribution for the target temperature. The Verlet algorithm[10] was used to integrate the equations of motion, with an integration time step of 1 fs (one femtosecond; 0.001 ps). To remove the high-energy bond-stretching contributions, the bond lengths were kept fixed at constant values during the MD runs, using the SHAKE procedure.[11] In order to reduce conformational changes occurring during MD simulations and resulting from end effects, the two terminal base pairs at each end were constrained to their minimized coordinates using the BELLY option of AMBER. Except for hydrogen bond lifetimes (see below), 200 structures were saved for each MD trajectory (one structure every 0.25 ps). The results of molecular dynamics were visualized with the program MDNM[12] on a graphic station PS330 Evans and Sutherland.

A recent analysis of the backbone behavior and of sugar puckering of different poly(dA)·poly(dT) models was done using Raman spectroscopy, molecular mechanics computations, and molecular dynamics simulations.[13] For those calculations, four models were studied: the model proposed by Lipanov and Chuprina,[14] model I; the model of Aymami et al.,[15] model II; the model suggested by Park et al.,[16] model III; and the heteronomous model, model IV, proposed by Arnott et al.[17] MD runs were performed on models I, III, and IV. After dynamical computations, a similar behavior is observed for all models. In terms of sugar puckering, the important result is the preference for the $C_{2'}$-endo domain of the adenine residues and for a conformation near $O_{4'}$-endo of the thymine residues. In that previous study,[13] it was concluded that best agreement between experimental and theoretical data is obtained with the sigmoidal distance-dependent dielectric function. Recent works[18] concluded that a sigmoidal distance-dependent dielectric function led to better agreement between experiment[19] and counterion condensation theory[20] than a constant-dielectric model.

Here, we pursue our analysis of MD simulations of models I and III. The nomenclature used for the homopolymers studied is as follows:



Existence Criteria for Three-Center Hydrogen Bonds

It was pointed out that the term "three-center" hydrogen bond is preferable to the use of "bifurcated" hydrogen bond in order to distinguish structures **1** from **2**.[21] The term "bifurcated"

hydrogen bond is used for case **2** where two hydrogen atoms are covalently bonded to a common donor atom X and hydrogen bonded to a same acceptor atom A. In the case of "three-center" hydrogen bond configuration (case **1**), the position of the hydrogen atom is determined by three nearest neighbor atoms, one donor atom to which it is covalently linked (X) and two acceptor atoms (A_1 and A_2).[21] In this work, the geometrical definitions for three-center hydrogen bonds as defined by Jeffrey[22] were used:

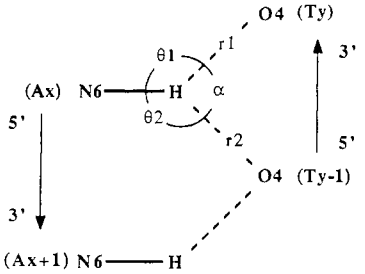
$$r_1 < r_2 < 3 \text{ Å}$$

$$\theta_1 > \theta_2 > 90°$$

$$350° < \theta_1 + \theta_2 + \alpha < 360°$$

According to Jeffrey,[22] three-center hydrogen bonds can be considered symmetrical if $r_1 \approx r_2$ and $\theta_1 \approx \theta_2$ and unsymmetrical if $r_2 - r_1 \approx 1$ Å. Symmetric three-center hydrogen bonds are less common.[23]

In poly(dA)·poly(dT), three-center hydrogen bonds occur between the adenine N6 amino group and thymine O4 atoms on two successive thymines:[5,24]



where r_1 refers to the distance between the hydrogen atom of N6(A_x) and O4(T_y) of the Watson–Crick hydrogen bond, and r_2 to the distance between the hydrogen atom of N6(A_x) and O4(T_{y-1}) of the three-center hydrogen bond. The Watson–Crick angle θ_1 refers to the (A_x) N6–H··O4(T_y) angle and the three-center angle θ_2 to (A_x)N6–H··O4(T_{y-1}). (T_y)O4··H(A_x)··O4(T_{y-1}) corresponds to the angle α.

The time dependence of the hydrogen bonds was evaluated in terms of mean bond lifetimes. Those were extracted from the simulation data in the following way. The "continuous" existence of a bond (i.e., the time elapsed before its first breakage) was assessed from a series of closely spaced structures saved during the MD simulations. The mean lifetime is then obtained by dividing the total time a H-bond exists by the number of times it is broken. With this type of analysis, short breakages followed by re-formation occurring between successive structures are not counted. Therefore, we used two measurement intervals (0.05 ps and 0.25 ps). Short lifetimes were smaller by a factor of 2 to 3 with the short interval than with the longer one. Longer lifetimes did not change appreciably. Because of the use of cutoff values, it is possible that the short lifetimes are underestimated.[25,26]

For calculations of autocorrelation functions, the history of each potential hydrogen bond was recorded as a series of 1 (if present) and 0 (if absent) defining the quantity $S(t)$.[25] The autocorrelation function is then given by

$$C(\tau) = \frac{\sum_{\tau_0=\tau_{min}}^{\tau_{max}} S(\tau_0)S(\tau_0 + \tau)}{\sum_{\tau_0=\tau_{min}}^{\tau_{max}} S(\tau_0)}$$

where $t_0 = \tau_0 \delta t$ is the time at which the measurement begins along the simulation run (with δt = 0.05 ps and $t_{min} = \tau_{min}\delta t$ = 5 ps). With this definition, bonds not formed at time t_0 are ignored, and,

(10) Verlet, L. *Phys. Rev.* **1967**, *159*, 98–103.
(11) Ryckaert, J. P.; Cicotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341. van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311–1327.
(12) Wurtz, J. M. Ph.D. Thesis, University Louis Pasteur, Strasbourg, France, 1988.
(13) Brahms, S.; Fritsch, V.; Brahms, J. G.; Westhof, E. *J. Mol. Biol.*, in press.
(14) Lipanov, A. A.; Chuprina, V. P. *Nucleic Acids Res.* **1987**, *15*, 5833–5844.
(15) Aymami, J.; Coll, M.; Frederick, C. A.; Wang, A. H. J.; Rich, A. *Nucleic Acids Res.* **1989**, *17*, 3229–3245.
(16) Park, H. S.; Arnott, S.; Chandrasekaran, R.; Milane, R. P.; Campagnari, F. *J. Mol. Biol.* **1987**, *197*, 513–523.
(17) Arnott, S.; Chandrasekaran, R.; Hall, I. H.; Puijaner, L. C. *Nucleic Acids Res.* **1983**, *11*, 4141–4155.
(18) Fenley, M. O.; Manning, G. S.; Olson, W. K. *Biopolymers* **1990**, *30*, 1191–1203. Fenley, M. O.; Manning, G. S.; Olson, W. K. *Biopolymers* **1990**, *30*, 1205–1213. Jayaram, B.; Swaminathan, S.; Beveridge, D. L.; Sharp, K.; Honig, B. *Macromolecules* **1990**, *23*, 3156–3165.
(19) Braulin, W. H.; Anderson, C. F.; Record, M. T., Jr. *Biochemistry* **1987**, *26*, 7724–7731. Padmanabhan, S.; Richey, B.; Anderson, C. F.; Record, M. T., Jr. *Biochemistry* **1988**, *27*, 4367–4376.
(20) Manning, G. *Q. Rev. Biophys.* **1978**, *11*, 179–246.
(21) Jeffrey, G. A.; Maluszynska, H. *Int. J. Biol. Macromol.* **1982**, *4*, 173–185.
(22) Jeffrey, G. A. *Landolt-Börnstein, Numerical Data and Functional relationships in Science and Technology, Group VII*, Volume 1, *Nucleic acids, Subvolume b, Crystallographic and Structural Data II*; Saenger, W., Ed. Springer-Verlag: Berlin-Heidelberg, 1989; pp 277–342.
(23) Taylor, R.; Kennard, O.; Versichel, W. *J. Am. Chem. Soc.* **1984**, *106*, 244–248.
(24) Yanagi, K.; Prive, G. G.; Dickerson, R. E. *J. Mol. Biol.* **1991**, *217*, 201–214. Westhof, E.; Beveridge, D. L. *Water Science Reviews*; Franks, F., Ed.; Cambridge University Press: London, 1990; Vol. 5, pp 24–136.
(25) Rapaport, D. C. *Mol. Phys.* **1983**, *50*, 1151–1162.
(26) Zichi, D. A.; Rossky, P. J. *J. Chem. Phys.* **1986**, *84*, 2814–2822.

Table I. Three-Center Hydrogen Bond Parameters (Distances (Å) and Angles (deg)) after Minimization and Dynamics for Model I and III as a Function of the Electrostatic Parameters[a]

		model I		model III		
		$\epsilon(r) = 4r$	$\epsilon(r) = \epsilon_{cal}$	$\epsilon(r) = 4r$	$\epsilon(r) = \epsilon_{cal}$	X-ray
$H\cdots O_4$	min	2.1	1.9	2.0	1.9	
(Watson–Crick)	dyn	2.4 (0.4)	2.0 (0.2)	2.4 (0.5)	2.0 (0.2)	1.9 (0.2)
	avg dyn	2.3	2.0	2.3	1.9	
$H\cdots O_4$	min	2.4	2.8	2.7	3.0	
(three-center)	dyn	2.6 (0.4)	2.7 (0.4)	2.8 (0.5)	2.9 (0.4)	2.8 (0.3)
	avg dyn	2.4	2.7	2.6	2.8	
$N_6\cdots O_4$	min	2.9	2.9	3.0	2.9	
(Watson–Crick)	dyn	3.2 (0.3)	2.9 (0.2)	3.3 (0.5)	2.9 (0.2)	2.9 (0.2)
	avg dyn	3.1	2.9	3.2	2.9	
$N_6\cdots O_4$	min	3.1	3.4	3.1	3.2	
(three-center)	dyn	3.1 (0.3)	3.2 (0.3)	3.2 (0.4)	3.2 (0.3)	3.2 (0.3)
	avg dyn	3.0	3.1	3.1	3.2	
$N_6\text{–}H\cdots O_4$	min	144	160	161	171	
(Watson–Crick)	dyn	142 (14)	153 (13)	146 (16)	158 (12)	152 (11)
	avg dyn	146	157	152	164	
$N_6\text{–}H\cdots O_4$	min	118	113	101	96	
(three-center)	dyn	118 (15)	109 (15)	111 (18)	102 (15)	104 (8)
	avg dyn	120	110	113	103	
$O_4\cdots H\cdots O_4$	min	94	87	92	90	
	dyn	90 (11)	91 (9)	88 (14)	92 (12)	96 (5)
	avg dyn	92	93	90	93	
sum of the angles	min	355	360	353	356	
	dyn	350 (12)	353 (9)	344 (19)	352 (11)	352 (8)
	avg dyn	357	360	356	360	

[a] The minimization and the dynamics values are averaged over the six base pairs allowed to move during MD simulations. The third line gives the values calculated with averaged dynamics structure (avg. dyn.). The X-ray column contains the averaged values obtained by Coll et al. and Nelson et al.,[5] resumed in the work of Yanagi et al.[24]

more importantly, bonds present at time *t*, whatever the number of intervening "breakage and re-formation" events, are included.

Results

Minimization. Table I gives the values of the three-center hydrogen-bond parameters, distance and angles averaged over the 10 base pairs. For all electrostatic parameters ($\epsilon(r) = 4r$ or $\epsilon(r) = \epsilon_{cal}$), the three-center hydrogen-bond distances are longer than the Watson–Crick ones. But those differences are more important with $\epsilon(r) = \epsilon_{cal}$ than with $\epsilon(r) = 4r$. Thus, with $\epsilon(r) = 4r$, the hydrogen bonds distribute themselves more symmetrically than with $\epsilon(r) = \epsilon_{cal}$. Whatever the electrostatic parameters or the model studied, our values are in agreement with the crystallographic values.[5,24] The "symmetrical" feature of the hydrogen bonds about the N6–H group obtained with $\epsilon(r) = 4r$ can also be observed in a smaller sum of angles than with $\epsilon(r) = \epsilon_{cal}$. Taylor et al.[23] have concluded that the out-of-plane departure of the hydrogen atom (here, deviation of the proton from the O4, N6, O4 plane) shows a small tendency to increase as the hydrogen bonds become more symmetrical so that the sum of angles decreases.

Table II illustrates some helical values describing the orientation of the central base pairs $A_4\text{–}T_{17}$, $A_5\text{–}T_{16}$, and $A_6\text{–}T_{15}$ before and after minimization. The near-zero values of the roll parameter stresses the tendency of base pairs to stack parallel to each other, whatever the model and the electrostatic parameters. High values of propeller twist are observed, but the magnitude depends on the model and on the dielectric function (−14° and −20°). But, whatever the homopolymer model, the magnitudes of the propeller twist values obtained with $\epsilon(r) = 4r$ are highest than those obtained with $\epsilon(r) = \epsilon_{cal}$. This feature agrees with the previously mentioned result: the three-center hydrogen-bonding system is more accentuated with $\epsilon(r) = 4r$ at the expense of the Watson–Crick H-bonds. Twist values are very similar independently of model and electrostatic parameters and near 36°.

Dynamics. Adenines A_1, A_2, A_9, and A_{10} and thymines T_{11}, T_{12}, T_{19}, and T_{20} were held fixed using the BELLY option of the program AMBER, while the remaining bases were allowed to move.

Table I gives the three-center hydrogen-bond parameters (distances and angles) averaged over the 200 saved structured for each MD simulation. Figures 1 and 2 describe these average values for each base pair of model I. Whatever the model studied,

Table II. Roll, Propeller Twist, and Twist Angle Values (in degrees) Averaged for the Three Base Pairs $A_4\text{–}T_{17}$, $A_5\text{–}T_{16}$, $A_6\text{–}T_{15}$ over 20 Structures[a]

		roll	propeller twist	twist
		Model I		
initial		−5	−12	36
minimization	$\epsilon(r) = 4r$	−1	−20	36
	$\epsilon(r) = \epsilon_{cal}$	−1	−15	36
dynamics	$\epsilon(r) = 4r$	−(1 ± 5)	−(24 ± 6)	36 ± 3
	$\epsilon(r) = \epsilon_{cal}$	−(2 ± 3)	−(19 ± 5)	36 ± 3
avg dynamics	$\epsilon(r) = 4r$	−2	−25	36
	$\epsilon(r) = \epsilon_{cal}$	−2	−20	37
		Model III		
initial		−3	−22	36
minimization	$\epsilon(r) = 4r$	−1	−18	37
	$\epsilon(r) = \epsilon_{cal}$	−2	−14	37
dynamics	$\epsilon(r) = 4r$	−(3 ± 5)	−(21 ± 8)	37 ± 3
	$\epsilon(r) = \epsilon_{cal}$	−(2 ± 5)	−(18 ± 5)	36 ± 3
avg dynamics	$\epsilon(r) = 4r$	−2	−24	38
	$\epsilon(r) = \epsilon_{cal}$	−2	−18	38
X-ray[5,24]		0	−20	36

[a] One structure saved energy 2.5 ps.

the differences between the Watson–Crick and the three-center H···O4 distances are more important with $\epsilon(r) = \epsilon_{cal}$ than with $\epsilon(r) = 4r$. Values obtained with $\epsilon(r) = \epsilon_{cal}$ agree better with the crystallographic ones. The use of $\epsilon(r) = 4r$ leads to an intermediate situation where the three-center hydrogen bonds are nearly as important as the Watson–Crick ones (Figure 1). If only the distances between N6 and O4 are considered, one would conclude that the three-center hydrogen bonds are more favorable than the Watson–Crick (Table I) (especially with $\epsilon(r) = 4r$ where the symmetrical situation, not found with $\epsilon(r) = \epsilon_{cal}$, is maintained during the whole MD simulation). Figure 3 illustrates these two situations: for model I with $\epsilon(r) = 4r$, the three-center hydrogen bonds, based only on distance criteria, occur nearly as often as the Watson–Crick ones, while with $\epsilon(r) = \epsilon_{cal}$ the Watson–Crick hydrogen bonds become preponderant. Another illustration of this phenomenon is given in Figure 4 representing the behaviors of two successive Watson–Crick hydrogen bonds with the intermediate three-center H-bond during a 50-ps MD simulation. With $\epsilon(r) = 4r$, the three curves are similar, while with $\epsilon(r) = \epsilon_{cal}$ the

Table III. Mean Lifetimes and Total Percentages of Existence in the MD Simulation Time for Each H-Bond Type[a]

	mean lifetimes in poly(dA)·poly(dT)							
	HN6(A)···O4(T)		N1(A)···HN3(T)		base pair		3-center	
	% (50 ps)	τ (ps)	% (50 ps)	τ (ps)	% (50 ps)	τ (ps)	% (50 ps)	τ (ps)
model I, $\epsilon = \epsilon_{cal}$	99.7	36	100.0	>50	100.0	>50	54–74	<1
model I, $\epsilon = 4r$	93.1	4	99.3	30	99.4	35	31–84	<2
model III, $\epsilon = \epsilon_{cal}$	99.9	46	100.0	>50	100.0	>50	43–55	<1
model III, $\epsilon = 4r$	89.1	6	97.1	37	97.5	41	26–68	<1

[a] For the three-center, the range indicates values obtained with strong and soft geometrical criteria (defined in legend of Figure 7).

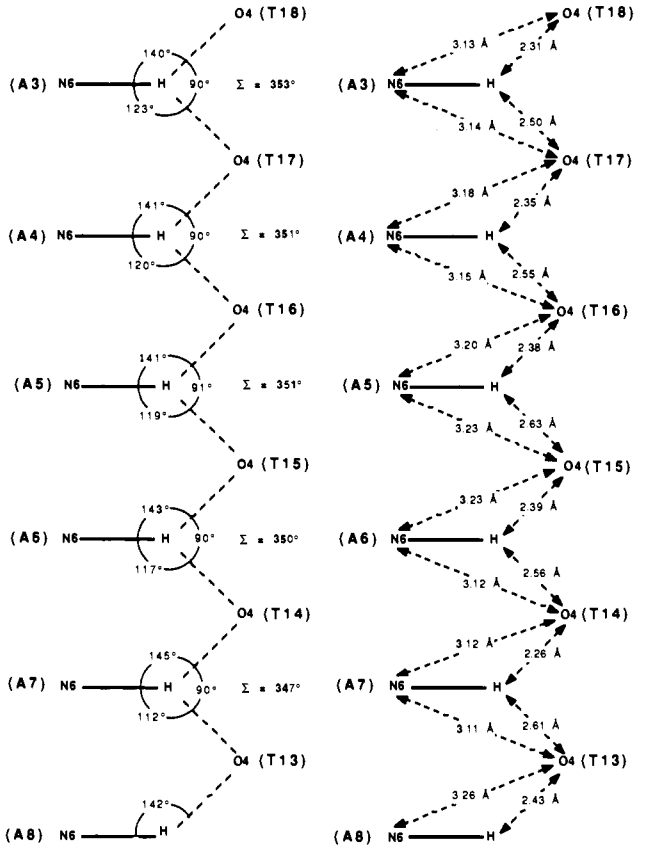


Figure 1. Three-center H-bond parameters (angles and distances) averaged over the 200 structures saved during a 50-ps MD simulation of model I with $\epsilon(r) = 4r$. The nomenclature used for the homopolymer is given in the text.

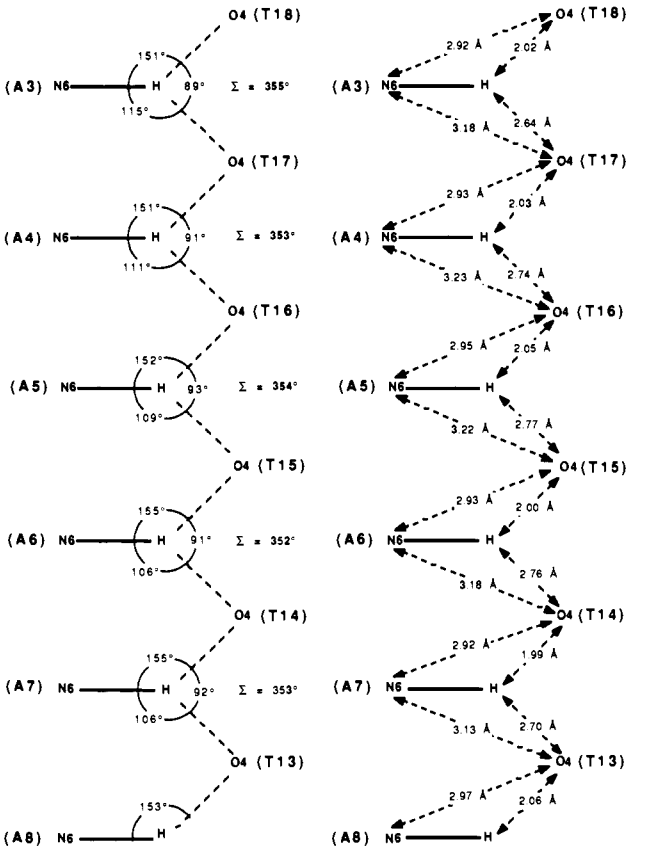


Figure 2. Three-center H-bond parameters (angles and distances) averaged over the 200 structures saved during a 50-ps MD simulation of model I with $\epsilon(r) = \epsilon_{cal}$. The nomenclature used for the homopolymer is given in the text.

curve for the three-center H-bond is near those obtained with $\epsilon(r) = 4r$ and different from those for the Watson–Crick H-bonds. Notice the importance of the hydrogen atom position; if only heavy atom positions are considered, such differential behaviors could not be observed.

With model III and $\epsilon(r) = 4r$, the Watson–Crick base pair A_7–T_{14} broke apart and stayed open about 7 ps before closing again (Figure 5). In this process, the A_7 adenine base moved into the major groove and the T_{14} thymine base made a strong three-center-type hydrogen bond with the A_6 adenine base.

The roll, propeller twist, and twist values, averaged over 20 structures saved during the MD simulations, are described in Table II. Concerning roll and twist values, one can see that they are very close to the minimized ones. A more important difference can be noticed for the propeller twist: for the two models and whatever the electrostatic parameters, the magnitude of propeller twist increased during the MD simulations (4–5°). The magnitude of the propeller twist value is more important with $\epsilon(r) = 4r$ than with $\epsilon(r) = \epsilon_{cal}$, in agreement with the hydrogen bond distance analysis which showed that the three-center hydrogen bond system is accentuated with $\epsilon(r) = 4r$. The comparisons between average values derived from X-ray structures[5,24] and from the simulations are excellent, but they do not allow a ranking of the dielectric functions.

Mean Lifetimes. To analyze the relative importance of the three-center hydrogen bonds, the mean lifetimes of the H-bonds were evaluated for the different models and electrostatic parameters. The values for the mean lifetimes of the two Watson–Crick hydrogen bonds and those of the three-center H-bonds involved in a A–T base pair are resumed in Table III. With $\epsilon(r) = \epsilon_{cal}$ and whatever the homopolymer model, the internal Watson–Crick bond N1(A)···HN3(T) occurs during the whole simulation; i.e., the geometrical criteria (distance smaller than 3 Å and angle greater than 90°) are satisfied during the whole 50-ps MD simulation. It is not the same for the external HN6(A)···O4(T) Watson–Crick hydrogen bond, which is broken within a 50-ps simulation. With $\epsilon(r) = 4r$, the differences observed previously are further amplified. The mean lifetimes of Watson–Crick hydrogen bonds are smaller than those obtained with $\epsilon(r) = \epsilon_{cal}$. For the HN6(A)···O4(T) bonds, lifetimes are seven to nine times longer with $\epsilon(r) = \epsilon_{cal}$. For the internal N1(A)···HN3(T), the differences are less important, but these hydrogen bonds are broken during a 50-ps MD simulation with $\epsilon(r) = 4r$, a situation which did not occur with $\epsilon(r) = \epsilon_{cal}$. It is known on the basis of NMR experiments[27] that the lifetimes of a base pair is of the order of

(27) Leroy, J. L.; Kochoyan, M.; Huynh-Dinh, T.; Gueron, M. *J. Mol. Biol.* **1987**, *200*, 223–238.

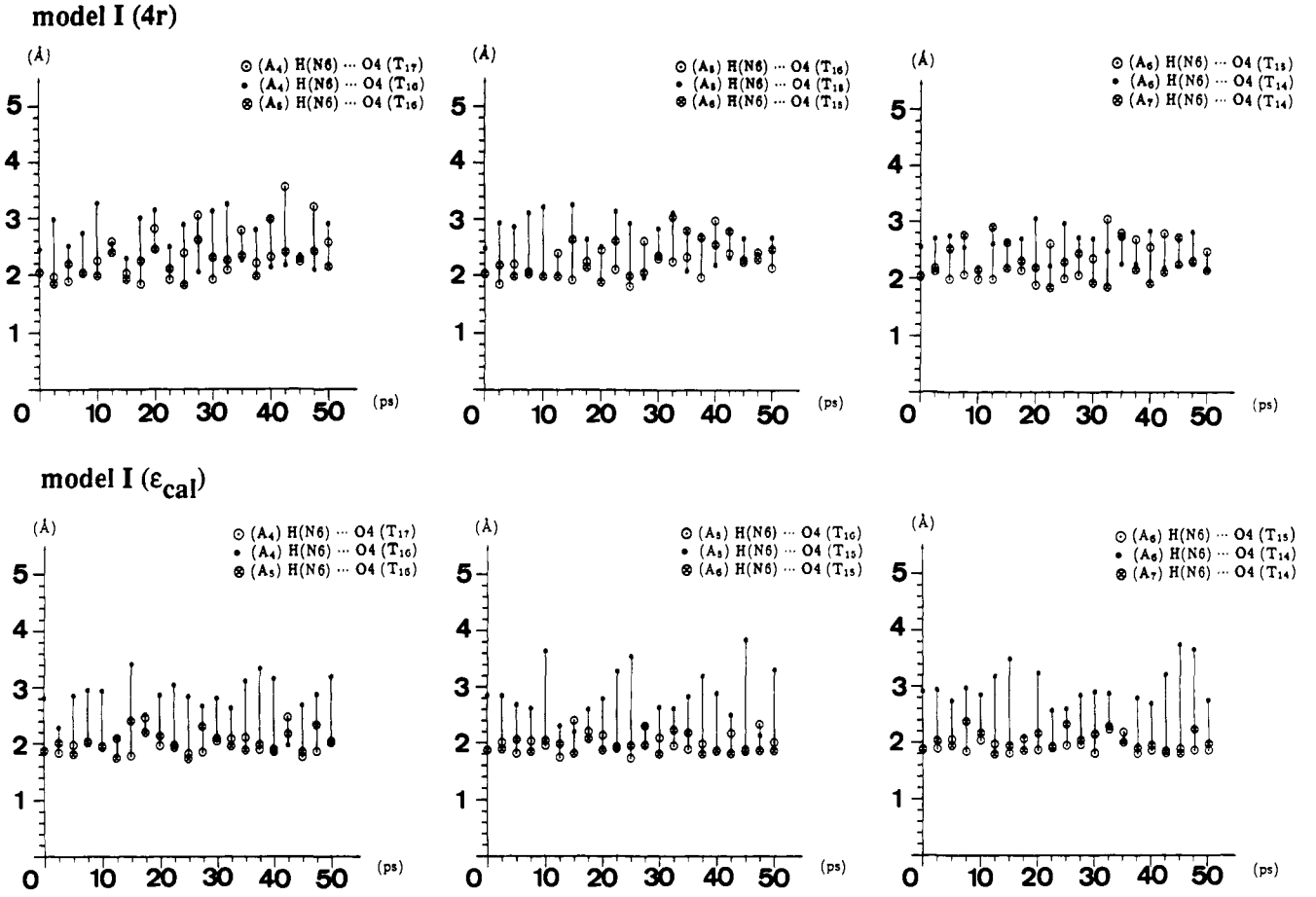


Figure 3. Variations of hydrogen-bond distances during the 50-ps MD simulation of model I with two dielectric functions. The dark point (●) corresponds to a three-center H-bond and the circles (⊙, ⊗) to the two Watson–Crick H-bonds around it. The three-center H-bonds observed are the following: HN6(A_4)···O4(T_{16}), HN6(A_5)···O4(T_{15}), and HN6(A_6)···O4(T_{14}).

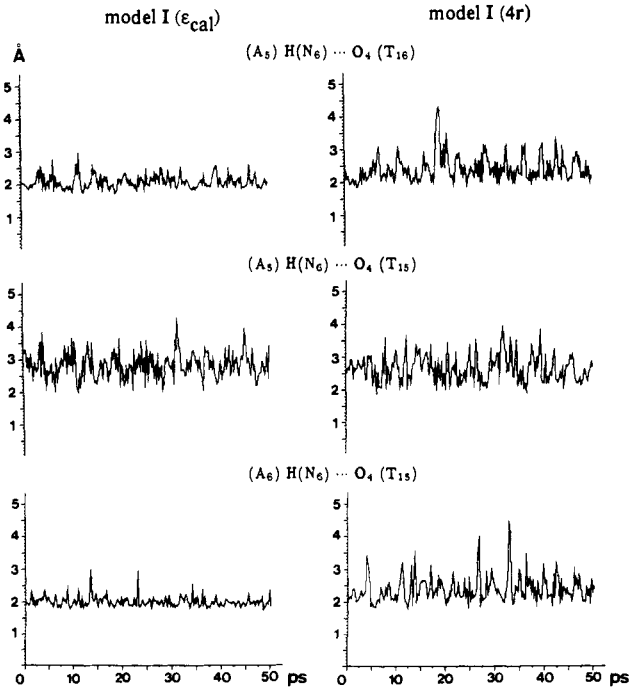


Figure 4. Another illustration of the differences in H-bond behaviors as a function of the electrostatic parameters. The curves illustrate the three-center H-bond HN6(A_5)···O4(T_{15}) behavior and the two Watson–Crick H-bonds around it.

1 ms (and up to 122 ms in stretches of A–Ts). At 300 K, with $\epsilon(r) = 4r$, the base pair is already broken during a 50-ps simulation, while it is not the case with $\epsilon(r) = \epsilon_{cal}$.

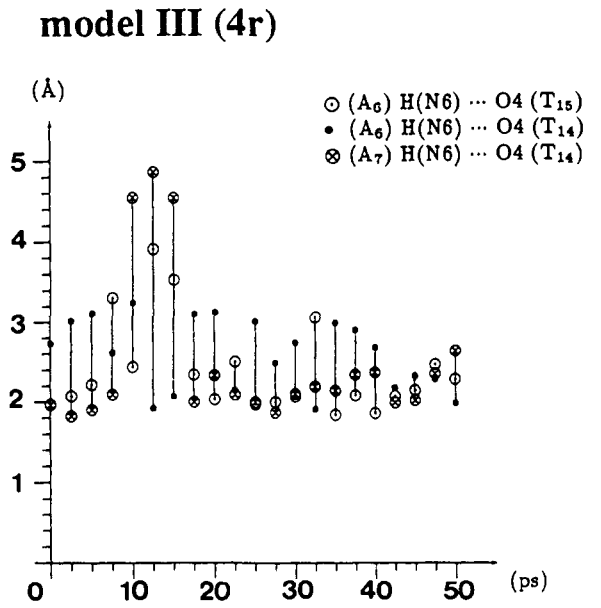


Figure 5. Evolution of the Watson–Crick base pair A_7–T_{14} (⊗) during a 50-ps MD simulation of model III with $\epsilon(r) = 4r$. During 7 ps (between 10 and 17 ps of the MD simulation), the A_7 adenine base stayed in the major groove with the T_{14} thymine base making a strong three-center H-bond with the A_6 adenine base (●).

The results of MD simulations performed at a series of temperature between 50 and 300 K are resumed in Table IV. For the computations between 50 and 200 K, the two Watson–Crick hydrogen bonds occur during the whole simulation. The external Watson–Crick hydrogen bond HN6(A)···O4(T) is broken within

Table IV. Mean Lifetimes and Percentages of Existence for Each H-Bond Type as a Function of the MD Simulation Temperature[a]

	mean lifetimes in poly(dA)·poly(dT) (model 1, $\epsilon = \epsilon_{cal}$)					
	HN6(A)···O4(T)		N1(A)···HN3(T)		3-center	
	% (50 ps)	τ (ps)	% (50 ps)	τ (ps)	% (50 ps)	τ (ps)
300 K	99.7	36.0	100.0	>50	54 (4)	0.7 (0.1)
	99.6	20.4	100.0	>50	51 (6)	0.2 (0.0)
250 K	99.9	45.8	100.0	>50	54 (5)	0.7 (0.1)
200 K	100.0	>50	100.0	>50	67 (4)	1.0 (0.3)
150 K	100.0	>50	100.0	>50	72 (7)	1.1 (0.3)
100 K	100.0	>50	100.0	>50	83 (6)	1.8 (0.7)
50 K	100.0	>50	100.0	>50	88 (6)	3 (2)
	100.0	>50	100.0	>50	88 (6)	1.0 (0.5)

[a] The measurement interval is 0.25 ps. For the three-center values, rms deviations are given in parentheses. At 50 K and 300 K, values are also given for a measurement interval of 0.05 ps (second line).

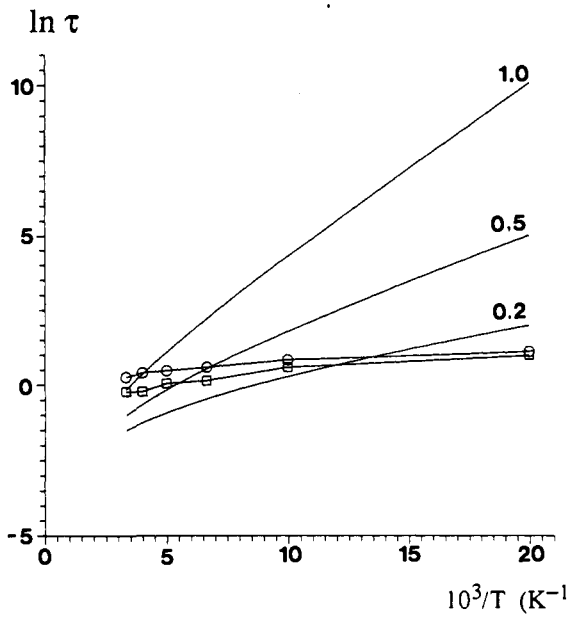


Figure 6. Plot of ln τ versus $1/T$ for the mean lifetimes of the three-center hydrogen bonds (□) and for the mean lifetimes of the adenine sugar in the $C_{2'}$-endo pseudorotational domain (○) for model 1, with ϵ_{cal}. The theoretical curves,[29] given by equation $\tau(ps) = \nu^{-1} \exp(\Delta G^*/RT)$, with $\nu^{-1} = 0.16$ ps, correspond to ΔG^* equal to 0.2, 0.5, and 1 kcal mol^{-1}.

the 50-ps simulation above 250 K. The mean lifetimes of the three-center hydrogen bonds increase significantly upon cooling; at 50 K, they are at least three times greater than at 300 K. But, although they are present about 88% of the simulation time at 50 K, the lifetime of the three-center hydrogen bonds is still on an average smaller by at least a factor of 20 than the lifetime of the Watson–Crick hydrogen bonds. Thus, the lifetimes of the three-center hydrogen bonds are of the same order as those of the hydrogen bonds in computer simulations of water.[26,28]

Figure 6 shows a plot of ln τ versus $1/T$ for the mean lifetimes of the three-center hydrogen bonds compared to the lifetimes of the adenine sugar in the $C_{2'}$-endo pseudorotational domain[13] for model I, with ϵ_{cal}. For comparisons, the theoretical curves expected on the basis of absolute rate theory[29] are given (with activation energies of 0.2, 0.5, and 1.0 kcal mol^{-1}). Clearly, the temperature dependence is much weaker and the curves for the two types of lifetimes are close to each other. Thus, we would conclude that the activation energies governing the three-center H-bonds are of the same order of magnitude as those governing the vibrational and pseudorotational movements in the puckered sugars. Very low activation energies (<1 kcal mol^{-1}) are observed also for

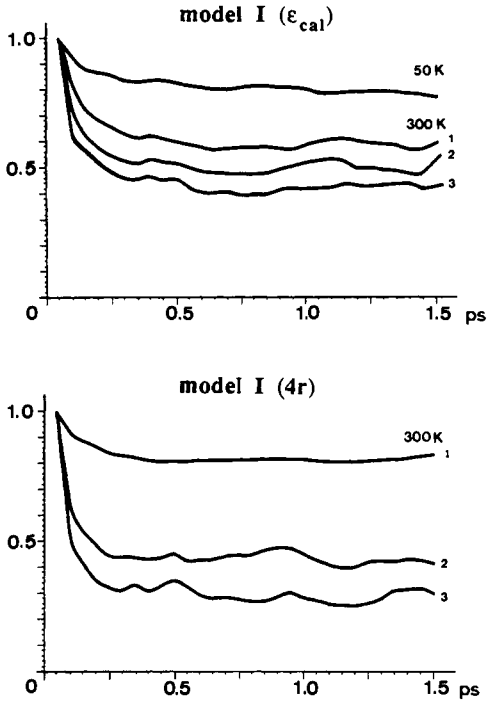


Figure 7. Autocorrelation functions for the three-center hydrogen bond between A_5 and T_{15}, with $\epsilon(r) = \epsilon_{cal}$ and $\epsilon(r) = 4r$ for (1) soft geometrical criteria ($r_2 < 3$ Å, $\theta_2 > 90°$); (2) medium geometrical criteria ($r_1 < r_2 < 3$ Å, $\theta_1 > \theta_2 > 90°$); and (3) strong geometrical criteria ($r_1 < r_2 < 3$ Å, $\theta_1 > \theta_2 > 90°$, $350° < \theta_1 + \theta_2 + \alpha < 360°$).

hydrogen-bond lifetimes in simulation of liquid water.[28]

In order to analyze further the dynamics of the hydrogen bonds, autocorrelation functions were evaluated. Two examples are shown in Figure 7. With $\epsilon(r) = 4r$, the autocorrelation curves depend strongly on the geometrical criteria used for defining the three-center hydrogen bonds. This is much less the case with $\epsilon(r) = \epsilon_{cal}$, indicating again the superiority of the ϵ_{cal} dielectric function. In all curves, there is a very rapid drop of autocorrelation followed by a smooth transition to a plateau value of 0.5, reached after 0.5 ps in the case of $\epsilon(r) = \epsilon_{cal}$. Thus, when observing a three-center H-bond, there is a 50% chance of observing it again after 0.5 ps. This value increases to 80% at low temperature.

Conclusions

Molecular dynamics simulations performed on several starting 3D models of (A,T) polymers lead to the conclusion that adenine residues fluctuate preferentially in the $C_{2'}$-endo domain while thymine residues occupy a domain closer to the $O_{4'}$-endo.[13] At the same time, the values for the glycosyl torsion angle, governing the orientation of the base with respect to sugar, are systematically higher in adenine than in thymine residues. These preferences induce double helical structures with important propeller twist values. This is especially the case of homopolymers where such a geometrical arrangement allows the formation of three-center hydrogen bonds. The present analysis of molecular dynamics simulations of dA·dT oligomers indicates that the three-center hydrogen bonds appear as geometrical consequences of the unusual structural properties of such oligomers, rather than as a factor contributing significantly to their stabilization. Indeed, the mean lifetimes and activation energies for the three-center hydrogen bonds are of the same order of magnitude as those for the hydrogen bonds in simulations of liquid water. However, the autocorrelation functions tend to an average value of 0.5 after 0.5 ps, indicating a high probability of occurrence. A similar conclusion was also reached after stereochemical refinement.[30]

The geometrical and chemical environment of the minor groove of dA·dT tracts is such that water molecules are able to bridge

(28) Geiger, A.; Mausbach, P.; Schnitker, J.; Blumberg, R. L.; Stanley, H. E. *J. Phys., Colloq. C7* **1984**, *45*, 13–30.

(29) Glasstone, S.; Laidler, K. J.; Eyring, H. *The Theory of Rate Processes*; McGraw-Hill: New York, 1941.

(30) Lipanov, A. A.; Chuprina, V. P.; Alexeev, D. G.; Skuratovskii, I. Ya. *J. Biomol. Struct. Dyn.* **1990**, *7*, 811–826.

hydration sites of adjacent residues on the two strands. Because of helical periodicity, such a situation leads to a regular hydration network called the hydration spine.[31] It is expected that this hydration spin contributes in an important way to the stabilization of the unusual properties of dA·dT polymers, and the evaluation of that factor will be the aim of further studies.

(31) Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1981**, *151*, 535-556. Chuprina, V. P. *Nucleic Acids Res.* **1987**, *15*, 293-311. Westhof, E. *Annu. Rev. Biophys. Chem.* **1988**, *17*, 125-144.

Acknowledgment. We thank G. S. Manning and D. L. Beveridge for continuing discussions and helpful comments on the theoretical treatment of the ionic environment on the modelling of nucleic acids. We also thank P. Koehl (IBMC, Strasbourg) and G. Wipff (Université Louis Pasteur, Strasbourg) for computer programs, and C. Gross, M. F. Janot, B. Speckel, and C. Tugene at the Centre de Calcul (Strasbourg-Cronenbourg) for help with the IBM 3090 system.

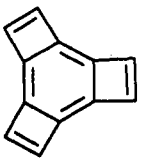
Registry No. Poly(dA)·poly(dT), 24939-09-1.

Is the Mills-Nixon Effect Real?

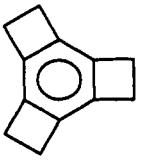
Amnon Stanger

Contribution from the Department of Chemistry, Technion-Israel Institute of Technology, Haifa 32000, Israel. *Received January 25, 1991. Revised Manuscript Received June 25, 1991*

Abstract: The long debated issue of the Mills-Nixon effect is reexamined. If benzene's hydrogens are bent as if to form small rings, pure strain is imposed on the aromatic system. The result of this strain is fixation of benzene's bonds according to the Mills-Nixon postulation. It was found (eq 1) that ΔR is proportional to $\sin^2 \alpha$, where ΔR is the difference between the long and the short bonds in the alternated benzene and α is the deviation from the natural bond angle (120°). However, the 3-21G-optimized geometries for **2** and **3** show that these systems are much more delocalized (i.e., ΔR is much smaller) than



2



3

is expected from the bent benzene model. The "effective" bond angles (i.e., the angles that should have caused the observed ΔR) of these molecules were calculated by using the ΔR values from the 3-21G-optimized geometries and the relationship found between the bending angle and ΔR (eq 1). Thus, 3-21G angles are 86.2 and 93.6° for **2** and **3**, respectively, whereas the "effective" angles (eq 1) are 92.0 and 108.1°, respectively. This difference was attributed to the formation of "banana bonds" in these systems. Comparison of the 3-21G results and the effective bond angles with the literature X-ray structures and electron-density-deformation analyses for benzenes annelated to three- and four-membered rings shows excellent agreement. It is concluded that when strain is imposed on a benzene the system responds by localizing its bonds in the Mills-Nixon manner. However, when small rings are annelated, the Mills-Nixon effect diminishes due to the formation of "banana bonds". Thus, special conditions must be fulfilled in order to observe the effect. Some predictions as to the geometries and electron densities of yet unstudied systems are given.

Introduction

The Mills-Nixon effect is a long debated issue.[1] Originally, it was suggested to explain reactivities and selectivities of benzenes annelated to small rings.[1a] However, during the years the subject evolved, and today it is known as the effect causing an aromatic moiety to localize its bonds (e.g., alternating arrangement of single and double bonds instead of the usual symmetric arrangement) due to strain imposed by small annelated ring(s) and hence to change the system's structure and reactivity. Some chemists claim that the effect is real[1c-f,i-k] whereas others suggest that it is an artifact of theoretical approximations, and higher level calculations show that the effect is not real.[1b,h,g] Experimentally, no evidence that unambiguously demonstrates the effect has been found, although it was looked for.[2]

Our interest in the issue began with the synthesis and structure determination of **1** by Diercks and Vollhardt.[3] The central ring of the molecule shows pronounced bond fixation (i.e., alternating C-C bond lengths of ca. 1.494 and 1.335 Å), whereas the outer rings are almost completely delocalized. The authors suggested that the reason for the bond localization in the central ring is the aromaticity-antiaromaticity interplay. The delocalization of the central ring would cause strong cyclobutadienic (antiaromatic) character in the four-membered rings. Thus, the authors assumed that the loss of aromaticity in one ring is energetically favorable as compared to the formation of three cyclobutadienes. On the other hand, the three four-membered rings form small angles with

(1) (a) Mills, W. H.; Nixon, I. G. *J. Chem. Soc.* **1930**, 2510. (b) Longust-Higgins, H. L.; Coulson, C. A. *Trans. Faraday Soc.* **1946**, *42*, 756. (c) Chung, C. S.; Cooper, M. A.; Manatt, S. L. *Tetrahedron* **1971**, *27*, 701. (d) Halton, B.; Halton, M. P. *Tetrahedron* **1973**, *29*, 1717. (e) Mahanti, M. K. *Indian J. Chem.* **1980**, *19B*, 149. (f) Hiberty, P. C.; Ohanessian, G.; Delbecq, F. *J. Am. Chem. Soc.* **1985**, *107*, 3095. (g) Apeloig, Y.; Arad, D.; Halton, B.; Randell, C. J. *J. Am. Chem. Soc.* **1986**, *108*, 4932. (h) Apeloig, Y.; Arad, D. *J. Am. Chem. Soc.* **1986**, *108*, 3241. (i) Dewar, M. J. S.; Holloway, U. K. *J. Chem. Soc., Chem. Commun.* **1984**, 1188. This paper presents MNDO calculations of tricyclopropabenzene (**6**) and finds a strong Mills-Nixon effect. However, MNDO tends to overemphesize bond fixation, as is evident from our calculations of **2** and **3** (unpublished). See also ref 1h. (j) Eckert-Maksič, M.; Hodošček, M.; Kovaček, D.; Mitič, D.; Maksič, Z. B.; Poljanec, K. *J. Mol. Struct.* **1990**, *206*, 89 and references therein. (k) Maksič, Z. B. Manuscripts in preparation.

(2) For example: Mitchell, R. H.; Slowey, P. D.; Kamada, T.; Williams, R. V.; Garratt, P. J. *J. Am. Chem. Soc.* **1984**, *106*, 2431 and ref 4 therein.

(3) Diercks, R.; Vollhardt, K. P. C. *J. Am. Chem. Soc.* **1986**, *108*, 3150.

0002-7863/91/1513-8277$02.50/0