Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 6537–6540, October 1985
Biophysics

A molecular dynamics simulation of double-helical B-DNA including counterions and water

(DNA hydration/DNA dynamics/DNA structure/DNA sugar pucker/electrostatic interactions)

G. L. SEIBEL, U. C. SINGH, AND P. A. KOLLMAN

Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143

Communicated by Robert G. Parr, May 20, 1985

ABSTRACT **We present the results of an atomic level molecular dynamical simulation of a 5-base-pair fragment of double-helical DNA with inclusion of water and sodium counterions and a complete description of their electrostatic interactions. The shape of the double helix is preserved throughout the simulation, and the helix repeat is calculated to be 10.0, in reasonable agreement with experimental results. The most flexible conformational angles in the structure are the glycosidic angle and the sugar pucker.**

Molecular dynamics simulations have the capability of providing a tremendous amount of useful information about the structural and dynamic properties of macromolecules and, in favorable cases, these structural features can be related to biological properties. Such simulations have had extensive application to proteins but relatively fewer applications to nucleic acids.

The simple energy functions one must use for such simulations are probably inherently more accurate for globular proteins with compensating charges than for the polyanionic nucleic acids with less spherical shapes. The difficulties in accurately representing the electrostatic effects in polyanionic nucleic acids are great, and each of the previous molecular dynamics studies of DNA has used a different approach: Levitt (1) removed all charges entirely, with the assumption that the water/counterion atmosphere would greatly damp them out; Tidor *et al.* (2) reduced the phosphate charge to −0.2 and used a distance-dependent dielectric constant; Singh *et al.* (3) used a fully anionic DNA with large "hydrated" counterions and a distance-dependent dielectric constant. In the latter two cases, the distance-dependent dielectric constant was used to mimic the effect of the solvent, H_2O, not explicitly included in the simulation.

Water was not included in the above calculations because of the tremendous computational overhead rather than anything more fundamental. Nonetheless, the question remained: what would be the results of a molecular dynamics simulation including H_2O, Na^+ counterions, and using a dielectric constant $\varepsilon = 1$; i.e., allowing "long-range" electrostatic effects? Furthermore, how would these results compare with a parallel simulation using one of the above simpler models and not including H_2O (3)? The need for such a test was heightened by Levitt's (1) observation that full inclusion of charges with $\varepsilon = 1$ but without water caused a severe disruption of the DNA helix.

Given an opportunity to work on a Cray made it possible to consider doing such a full simulation of d(CGCGA)·d(TCGCG) with inclusion of counterions and H_2O and using $\varepsilon = 1$. The conversion of the molecular dynamics module of AMBER (4) from the VAX to the Cray X-MP was straightforward, as it was written in standard FORTRAN (American National Standards Institution). The subroutine that evaluates the nonbonded interactions constitutes the rate-limiting step in such calculations, so it was rewritten to take maximum advantage of the vector architecture of the Cray. The performance increase on the Cray was substantial, ≈150 times the speed of a VAX 11/780. This represents a 3-fold increase in speed compared to empirical energy calculations on earlier Cray computers, and it is due to the gather/scatter hardware incorporated in the X-MP 4/8. No assembly language coding was required to achieve this gain. The entire simulation required 20 hr of Cray time, and it would have taken 5–6 months on a dedicated VAX.

Our model system was constructed by using the Arnott (5) geometry of d(CGCGA)·d(TCGCG), which is the terminal sequence of the dodecamer solved crystallographically by Dickerson and coworkers (6). Na^+ counterions were placed on the phosphate bisector 3 Å from the P atom, and the solute DNA was placed into a large water bath constructed of repeating cubes of TIPS3P H_2O molecules, which were a snapshot from a Monte Carlo simulation of liquid water (7). We then removed those waters that were sterically disallowed or were >9 Å from any solute atom, leaving a total of 830 water molecules. After 500 cycles of molecular mechanics refinement and 8 psec of molecular dynamics equilibration, we carried out a further 106 psec of molecular dynamics calculations using a nonbonded cutoff of 10 Å. The averages reported below are for these 106 psec. We did not use periodic boundary conditions here, because our goal was to examine the effect of water and counterions on the DNA structure rather than the H_2O properties *per se*. We also used a rather larger time step than is typical for such simulations, 0.002 psec. Nonetheless, the average temperature stayed near 298 K throughout the simulation, and no artifacts due to inaccuracies in the numerical integration were apparent. The force-field parameters used for the DNA atoms are in ref. 8.

During the simulation, no H_2O molecules "evaporated" and six of the eight Na^+ counterions stayed near the corresponding phosphates where they were initially placed. However, two neighboring sodium ions initially attached to the GpC and CpA of the first chain moved a significant distance from their starting position, the former ending up near the edge of the water (still strongly coordinated to H_2O molecules) and the latter moving into the minor groove of the DNA (Fig. 1). Given the lack of periodic boundary conditions in our simulations, it was surprising to see such large movements of the Na^+; these movements likely represent only a "tip of the iceberg" of Na^+ mobility one would observe in a more extensive "water bath."

All of the average dihedral angles in the DNA remain in the same range as found in our earlier simulation without the explicit water molecules, but the rms atomic motions from the initial (equilibrated) structure were in the range of 0.5 Å and thus 50–60% of those observed in the simulation without water (3). There are some interesting differences between the overall structure of the DNA in the two simulations. Table 1 compares the averages for helical repeat, base twist, and base

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

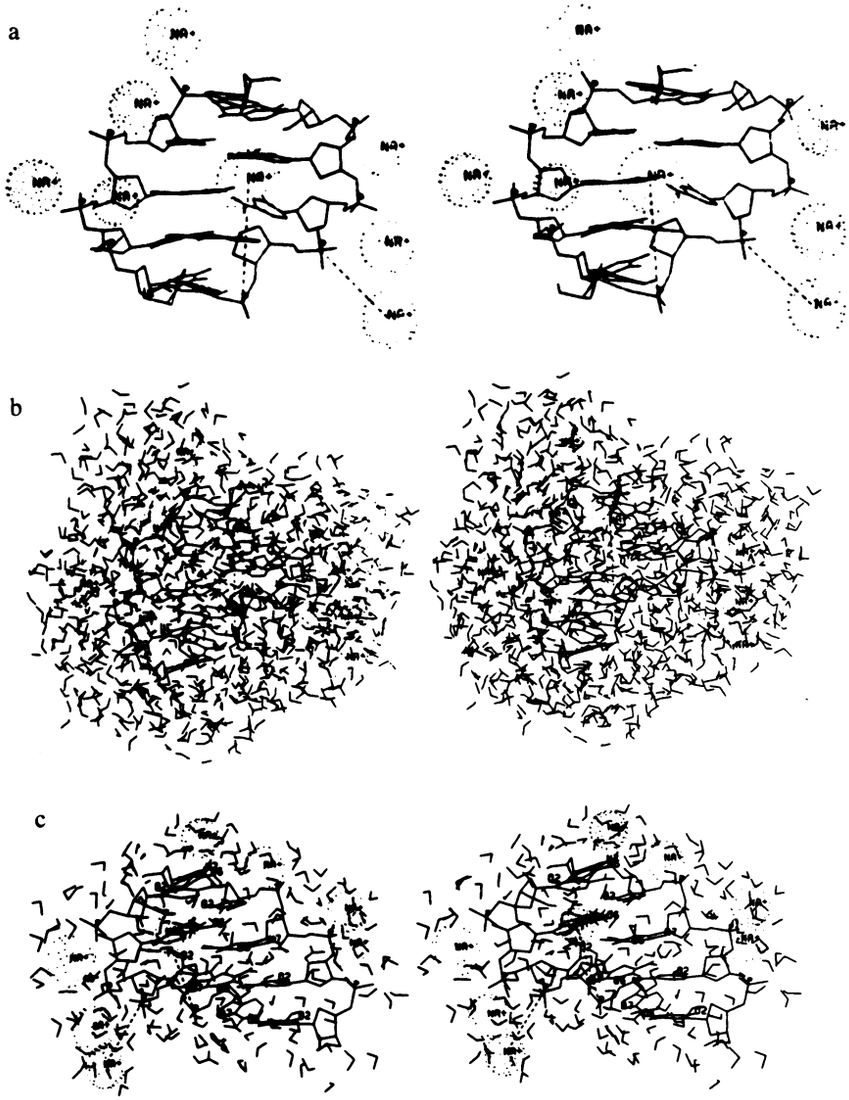

FIG. 1. (*a*) Stereoview of DNA and Na^+ at the end of simulation with Na^+ counterions shown as spheres. (*b*) Same as in *a* but with all the H_2O molecules included. (*c*) Same as in *a*, including only those H_2O within 3 Å of any DNA atom.

tilt in the two simulations. The trends for the helical repeat angles are similar, as is the average repeat angle for the entire structure (≈10.0). Nonetheless, there is considerably more tilt and twist for the central base pair in the presence of water. This is also reflected in the fact that the glycosidic angles for the central cytosine and guanine bases are considerably different from the rest of the bases. The phosphodiester angle ω' has a similar average in both simulations, but it has a much larger standard deviation (±20°) in the absence of water than in its presence (±12°), showing the role of water-phosphate hydrogen bonding in damping out large phosphate movements.

It is interesting that the sugar pucker profile is somewhat different in the two simulations (Table 2), in that many more examples of deviations from the classic C2′ *endo* conformation are found in the simulation with water than in the one without. NMR studies find a typical deoxyribo sugar to be 70–80% C2′ *endo*; 20–30% C3′ *endo* even in a double helix (11). Thus, our findings of a significant C3′ *endo* population and the absence of a significant percentage of conformations in the region with sugar pucker phase W from 200° to 340° are qualitatively consistent with observations, even though 100 psec is far too short to fully "equilibrate" the sugar conformations. The fact that the second sugar of the first chain remains in the O1′ *endo* region throughout the simulation, despite the fact that such a conformation is 1.5–2 kcal/mol (1 cal = 4.184 J) higher than C2′ *endo*, is interesting and is supported by the observation of a surprisingly large O1′ *endo* population in the crystal structure of Dickerson and coworkers (6). That structure is of low enough resolution that such "high-energy" puckers may be due to the x-ray refinement, but our calculations suggest that, given there are much stronger electrostatic forces on the DNA, changing the sugar pucker may likely be the least energetically costly way to compensate for them.

The Watson–Crick hydrogen bonds remain nearly intact throughout the simulation, although they are on average somewhat longer in the simulation with water than in the one without. In the simulation with water, however, the $NH_2 \cdots CO$ H-bond length between adenine-6 and thymine-4 is very long (2.58 ± 0.57 Å), whereas in the simulation without water, this H-bond length is 2.02 ± 0.16 Å. On the other hand, the adenine–thymine H-bond (A-N1 ··· H3-T) length remains at a much shorter distance of ≈1.9 Å in both simulations.

To focus on the effects that water might have on the DNA structure, we have analyzed in more detail the structure of the DNA–Na^+–H_2O interactions. For example, we have evaluated the average number of H_2O molecules within 3.1 Å of each Na^+ ion, as well as the angle the Na–O line makes with the H_2O bisector. An idealized C_{2v} symmetry Na^+–OH_2

Table 1. Helix parameters

Base (pair)*	Without H_2O†	X-ray‡	With H_2O§
	Twist		
C_1·G'_5	22.4 ± 8.6	27.1, 14.8	27.8 ± 13.4
G_2·C'_4	24.2 ± 11.0	22.1, 17.3	12.7 ± 6.0
C_3·G'_3	13.0 ± 7.6	6.3, 6.4	31.7 ± 9.3
G_4·C'_2	14.4 ± 6.6	22.2, 21.6	25.2 ± 7.5
A_5·T'_1	21.6 ± 11.3	24.2, 24.6	25.1 ± 10.9
Average	19.1	20.4, 16.9	24.5
	Tilt		
C_1	26.1 ± 11.5	12.6, 37.6	26.7 ± 10.8
G_2	12.1 ± 7.1	6.4, 9.1	10.9 ± 5.5
C_3	11.6 ± 5.6	7.5, 8.5	20.0 ± 8.3
G_4	10.5 ± 5.5	8.9, 10.9	17.2 ± 3.9
A_5	13.1 ± 6.5	16.0, 15.3	16.2 ± 5.6
T'_1	21.1 ± 8.0	9.2, 9.0	27.1 ± 10.7
C'_2	18.5 ± 10.4	12.7, 13.6	21.6 ± 4.8
G'_3	8.7 ± 4.5	1.1, 2.2	17.2 ± 3.7
C'_4	9.1 ± 4.7	20.8, 22.1	12.9 ± 5.9
G'_5	10.8 ± 7.0	4.8, 11.7	10.8 ± 8.1
Average	14.2	10.0, 14.0	18.0
	Helix repeat		
$C_1G'_5$·$G_2C'_4$	36.5 ± 4.5	35.0, 36.6	36.9 ± 2.7
$G_2C'_4$·$C_3G'_3$	34.8 ± 5.7	41.7, 39.3	27.9 ± 2.3
$C_3G'_3$·$G_4C'_2$	38.8 ± 6.8	28.7, 30.4	47.6 ± 3.2
$G_4C'_2$·$A_5T'_1$	33.8 ± 4.7	42.3, 38.0	32.3 ± 1.9
Average	36.0	36.9, 36.1	36.2

Twist is designated by the average angle the base pair planes make with each other (in degrees). Tilt is designated by the angle the bases make with the helix axis defined by the phosphate group (in degrees). Helix repeat is designated by the angle made by the successive N1(N9)–N1(N9) vectors projected onto the helix axis. Subscripts indicated position of the base.

*The bases of the first strand are CGCGA; the primed strand bases are TCGCG.

†Average values for TWIST, TILT, and HELIX REPEAT in the simulation with large "hydrated counterions" and no water; see ref. 3.

‡Values found in x-ray structure; ref. 6.

§Values found in this study.

Table 2. Sugar pucker profiles

Base*	q†	% of time sugar spent with various pseudorotation phases W‡				
		C3′ *endo*	O1′ *endo*	C2′ *endo*	O1′ *exo*	C1′ *endo*
C_1	0.34 ± 0.06	0.00	12.74	86.13	1.13	0
G_2	0.37 ± 0.05	6.32	91.79	1.89	0	0
C_3	0.42 ± 0.04	0.00	0.00	100.00	0	0
G_4	0.39 ± 0.05	98.96	0.00	0.00	0	1.04
A_5	0.34 ± 0.07	0.00	0.00	96.51	3.49	0
T'_1	0.34 ± 0.06	20.38	67.45	12.17	0	0
C'_2	0.36 ± 0.06	0.38	71.79	27.83	0	0
G'_3	0.39 ± 0.06	79.62	6.70	13.68	0	0
C'_4	0.31 ± 0.07	3.68	2.54	93.11	0.75	0
G'_5	0.38 ± 0.06	0.00	0.19	99.25	0.57	0

*Base to which sugar is attached on first strand is d(CGCGA); on second strand (with ′ notation), it is d(TCGCG). Subscripts indicate position of base.

†Average out-of-plane distance of sugar ring averaged over the 106-psec run (see ref. 9).

‡See ref. 10. C3′ *endo*, W = −18° to 54°, O1′ *endo*, W = 54°–126°; C2′ *endo*, W = 126°–198°; O1′ *exo*, W = 198°–270°, and C1′ *endo*, W = 270°–342°.

interaction should have this angle at 180°, but any angle greater than ≈100° should be an attractive interaction. For each of the eight sodium ions the average number of waters that have R(O–Na^+) < 3.1 Å and the Na^+–O bisector angle >100° was calculated and their average R(O–Na^+) distance was determined. For the six Na^+ that remain closely coordinated to the phosphate, R(Na^+–O) = 2.5–2.6 Å and the coordination numbers range from 4.6 to 5.1 and average 4.8; for those two Na^+ that move from the corresponding phosphate, R(Na^+–O) ≈ 2.5–2.6 Å and the average coordination number is 6.0.

The waters also form an extensive hydrogen-bonded network with the hydrophilic (X) atoms on the DNA backbone and major and minor grooves. If water is a proton donor, a linear H bond would correspond to an angle between the X–O line and OH_2 bisector of ≈50°, so we used a criterion of an X–O distance of <3.5 Å and X–O bisector angle between 0° and 80° as a criterion for H bonding. With such a criterion, the anionic phosphate oxygens OA and OB form an average of 3.0 H bonds (from 2.2 to 3.9 for the 16 such atoms), each with water molecules, with average distance 2.7–2.8 Å; the sugar O1′ average 1.1 H bonds ≈3.1 Å long, O3′ and O5′ (excluding the terminal O3′ and O5′) average 1.7 and 0.9 H bonds, respectively, with an average length of 3.1–3.2 Å. The smaller number of H bonds formed by O5′ than O3′ is likely due to the fact that O5′, when the ψ angle is gg (as it remains throughout the simulation), is "over" the sugar ring and more shielded from solvent than is O3′. It is clear, however, that backbone H bonding is dominated by the phosphate anionic oxygens, OA and OB, whose coordination properties are quite similar to those found in a Monte Carlo simulation of dimethyl phosphate (12).

There is extensive H bonding to the pyrimidine carbonyl O2, the purine N3, and the exposed guanine NH (HN2B) in the minor groove, with an average of 1.0 H bond and an O–O distance of 2.8–2.9 Å to O2, 0.9 H bonds and an O–N distance of 3.1 Å to N3, and 0.7 H bonds and an average O–N distance of 2.9–3.0 Å for the guanine NH_2 ... OH_2 interaction. These averages exclude the thymine O2 on the end, which is more exposed and forms, on average, 2.1 H bonds.

The carbonyl groups (guanine O6, thymine O4) in the major groove form more H bonds (1.4 for the central three, 2.7 for the outer two; R_{av}, ≈2.9 Å) than those in the minor groove, with R_{av} = 1.9 Å and N_{av} = 1.6 for the exposed H of the cytosine $4NH_2$ and adenine $6NH_2$ and R_{av} = 2.9 Å, N_{av} = 1.6 for the purine N7.

There are no "close" (i.e., shorter than the sum of van der Waals radii) contacts involving the waters and the hydrophobic atoms—C1′, C2′, C3′, C4′, and C5′ of the sugars, C8 (purines), and C5, C6, and C7 (pyrimidines). Our Monte Carlo simulations (12) on dimethyl phosphate in water reveal that the water structure near the dimethyl phosphate methyl groups is similar to the nature of water around pure hydrocarbons, in which water molecules tend to stick their lone pairs and hydrogen atoms away from the hydrophobic atoms. It is likely that this physical effect is operative here also, but the fact that the hydrophobic and hydrophilic atoms in DNA are intermingled in a complex way precludes a simple analysis.

All in all, the nature of water–DNA and water–Na^+ interactions observed in these simulations are quite consistent with chemical intuition, x-ray structures (13), and previous Monte Carlo (14) and molecular dynamics calculations (10) on rigid DNA–Na^+–H_2O interactions: extensive and strong H_2O–Na^+ and OH_2–anion phosphate oxygen interactions, weaker but still observable H bonds to the hydrophilic atoms on the backbone O5′, O3′, O1′, major grooves O6, HN6A, N7, HN4A, O4 and minor grooves O2, N3, and HN2B. These latter H bonds are expected to be comparable in strength to water–water H bonds. Thus, even though the focus of this work has been on the nature of DNA flexibility in the presence of counterions and water rather than the details of DNA–Na^+–H_2O interactions, the fact that the

nature of these H_2O interactions is consistent with one's intuition adds further credibility to this study, particularly given the fact that no periodic boundary conditions were used here.

It is interesting that the sequence we have studied, d(CGCGA)·d(TCGCG), would be expected to be near the single- to double-stranded transition temperature (dependent on the ionic strength of the solution) at 25°C. Thus, the fact that we observe significantly greater average base tilt, twists, and longer Watson–Crick H bonds in the simulation with H_2O than in the simulation without H_2O is consistent with the inherent instability of such a short double-helical structure in aqueous solution. Nonetheless, $\approx 10^{-10}$ sec is certainly too short to observe such a transition. We note, however, that despite this inherent instability, our helix apparently remains more B-DNA-like than Levitt's simulation (1) on a longer more inherently stable double helix. This suggests that further molecular simulations with full inclusion of water, counterions, and a realistic representation of DNA charges are likely to be a useful complement to experiments on DNA properties. In addition, despite the fact that there were many differences in details between our two simulations, one with full inclusion of H_2O, counterions, and $\varepsilon = 1$ and one with a simpler representation with no explicit H_2O, the qualitative agreement in helix repeat, average dihedral angle, and overall shape of the double helix suggests that the simpler model is certainly worth further exploration as well, on other and longer DNA sequences and different DNA structures (A, B, and Z). In fact, we already have done so on $dA_{10}\cdot dT_{10}$ and $d(ATATATATAT)_2$ and have found a significantly smaller helix repeat in the homopolymer $dA_{10}\cdot dT_{10}$, consistent with experimental results (unpublished).

We are grateful to John Aldag and CRAY Research (Mendota Heights, MN) for the computer time provided and for a stimulating summer of 1984 spent by one of us (G.L.S.). We thank W. van Gunsteren, who provided the molecular dynamics code that was incorporated into AMBER by U.C.S., M. Levitt for the DNA display program that enabled a movie to be made of this simulation, the National Institutes of Health for grant support to P.A.K. (CA-25644) and the UCSF Computer Graphics Lab (supported by RR-1081, P. I. R. Langridge, system manager, T. Ferrin) for the excellent display facilities.

1. Levitt, M. (1983) *Cold Spring Harbor Symp. Quant. Biol.* **47,** 271–275.
2. Tidor, B., Brooks, B. & Karplus, M. (1983) *J. Biomol. Struct. Dyn.* **1,** 231–252.
3. Singh, U. C., Weiner, S. J. & Kollman, P. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 755–759.
4. Weiner, P. & Kollman, P. (1981) *J. Comp. Chem.* **2,** 287–303.
5. Arnott, S., Campbell-Smith, P. & Chandresekharan, P. (1976) *CRC Handbook of Biochemistry* (CRC, Boca Raton, FL), Vol. 2, pp. 411–422.
6. Drew, H., Wing, R., Takano, T., Broka, E., Tanaka, S., Itakura, K. & Dickerson, R. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 2179–2183.
7. Jorgensen, W., Chandrasekhar, J., Madura, J., Impey, M. & Klein, R. (1983) *J. Chem. Phys.* **79,** 926–935.
8. Weiner, S., Kollman, P., Case, D., Singh, U., Ghio, C., Alagona, G., Profeta, S. & Weiner, P. (1984) *J. Am. Chem. Soc.* **106,** 765–784.
9. Cremer, D. & Pople, J. (1975) *J. Am. Chem. Soc.* **97,** 1354–1358.
10. Lee, W. K., Gao, Y. & Prohofsky, E. W. (1984) *Biopolymers* **23,** 257–270.
11. Mellema, J., Haasnoot, C., van der Marel, G., Wille, G., van Boeckel, C., van Boom, J. & Altona, C. (1983) *Nucleic Acids Res.* **11,** 5717–5738.
12. Alagona, G., Ghio, C. & Kollman, P. (1985) *J. Am. Chem. Soc.* **107,** 2229–2239.
13. Kopka, M., Fratini, A., Drew, H. & Dickerson, R. E. (1983) *J. Mol. Biol.* **163,** 129–146.
14. Clementi, E. & Corongiu, G. (1982) *Int. J. Quantum Chem.* **22,** 595–630.