INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
34100 TRIESTE (ITALY) - P.O.B. 586 - MIRAMARE - STRADA COSTIERA 11 - TELEPHONES: 224281/2/3/4/5/6
CABLE: CENTRATOM - TELEX 460392 - I

SMR/111 - 26

SECOND SUMMER COLLEGE IN BIOPHYSICS

30 July - 7 September 1984

Lecture 3: DNA and proflavin crystals: determination of water structure. Complementarity of diffraction experiments and simulation.

G. CORONGIU
IBM Corporation
IS/TG
Poughkeepsie, New York 12602
U.S.A.

These are preliminary lecture notes, intended only for distribution to participants. Missing or extra copies are available from Room 230.

Journal of Biomolecular Structure and Dynamics, ISSN 0739-1102
Volume 1 (1983), ©Adenine Press (1983).

Networks of Water Molecules in a Proflavine Deoxydinucleoside Phosphate Complex

Kwang S. Kim
National Foundation for Cancer Research
Poughkeepsie Laboratory

and

G. Corongiu and E. Clementi
IBM Corporation, IS & TG, Dept. D55
P.O. Box 390, Poughkeepsie, N.Y., 12602

Abstract

Using previously reported ab-initio atom-atom potentials for the interactions of a water molecule with phosphates, sugars and bases and newly computed ab-initio atom-atom potentials for the interaction between a proflavine ion and water, we have analyzed with the Monte-Carlo Metropolis method networks of water molecules hydrating a 2:2 complex of proflavine and deoxycytidylyl-3',5'-guanosine, recently studied with X-ray crystallography. From our simulations we have i) verified the quality of our atom-atom potentials by obtaining patterns of oxygen atoms in very good agreement with the X-ray patterns for the minor groove and in reasonable agreement in the major groove, ii) predicted the water's hydrogen atoms positions and iii) preliminarily predicted the number of water molecules not reported in the X-ray study but present in the major groove. The above data, even if preliminary, and the analyses on the energetics of the water-water, water-proflavine and water-dCpG interactions indicate that very detailed accounts on the water filaments in the above crystal can be obtained optimally by merging computer and X-ray experiments.

Introduction

Recently a number of X-ray studies have focused on the structure of drug-DNA complexes, in particular crystals of proflavine-oligonucleotide complexes. In general, these crystals contain a rather large number of water molecules per unit cell.[1,2,3] This feature is of particular interest in this laboratory because of our research program directed at elucidating the structure of water molecules, hydrating ions and biomolecules, in general,[4] nucleic acids in particular.[5,6,7]

In this work we shall analyze the network of water molecules contained in a 2:2 crystal of proflavine ion (PF) and deoxycytidylyl-3',5'-guanosine, (dCpG), comparing our results with those reported by S. Neidle, H. M. Berman and H. S. Shieh.[3]

This comparison will allow for a verification of some aspects of the theoretical-computational methods we have been advocating and, at the same time, for a verification of the water network derived in the X-ray study.[2,3] Concerning the latter we recall that from the X-ray diffraction data, one cannot obtain the water's hydrogen atoms positions (and hence the orientation of the water molecules) nor can one observe those water molecules which are too mobile. We recall in addition that in the reported crystal structure more than the reported 50 water molecules are believed to be present as one might infer from studies on a second crystal.[3,8]

Concerning the methodological verification, with this work we can test again the predictive power of our atom-atom pair potentials as used in simulations of large and complex systems. As known, these potentials are not as accurate and refined as those used in simpler chemical systems, where accurate two- and three-body quantum mechanical computations can be performed.[9]

In our Monte Carlo simulations the predicted water molecule's positions and orientations have been computed in the field of the dCpG and PF species; the latter are assumed as rigid. In our simulation the coordinates of the skeleton atoms for PF and dCpG are those of Ref. (2), to which we have added the corresponding hydrogen atoms, using standard bond length and bond angles. From the X-ray study we know that in the 2:2 complex, PF:dCpG, the two dinucleoside phosphate strands form self-complementary duplexes with Watson-Crick hydrogen-bonds. One proflavine is asymmetrically intercalated between the base pairs and the other is stacked above them; this is shown in Figure 1, projecting the complex into the x-y and into the x-z planes. We know from Ref. (2) that a single unit cell contains four complexes or four "asymmetric units" arranged in such a way as to define a major and a minor groove. In Figure 2 we report a x-y projection with 16 complexes (or 4 unit cells); the pattern shows "large" cavities separated by "small" ones. Since the system has translational symmetry along the z axis, these cavities are referred to as *major* and *minor grooves*, respectively. Water molecules fill in these grooves. In a crystal 50 water molecules have been found by X-ray in the two grooves corresponding to one half unit cell; in a second crystal 55 water molecules have been found. We recall that the length, width and height are approximately x=18Å, y=5Å, z=13.5Å for the major groove and x=9Å, y=3Å and z=13.5Å for the minor groove, respectively.

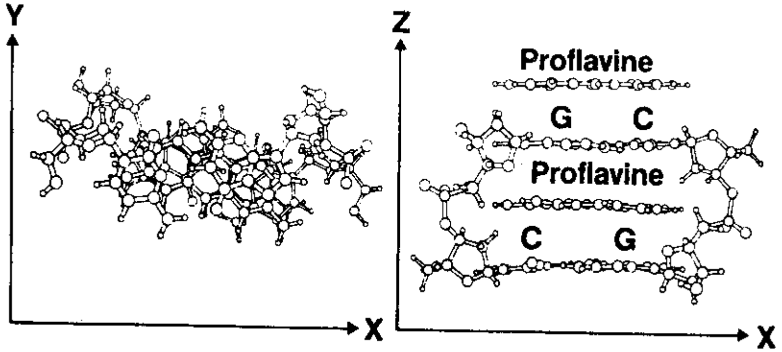

Figure 1. Asymmetric unit of (dCpG:PF+)2 projected into the x-y and x-z plane, respectively.

In our experiment the Monte Carlo simulation cannot be limited to a volume as small as a single unit cell as it is often done for neutral solutes. Indeed the presence of the negatively charged groups, PO_4^-, and of the positively charged proflavines, PF^+, *compels one to consider a much larger sample in order to take into account the long-range ionic field which extends further than the dimensions of the unit cell.* We have selected a sample which is three times as large as the one shown in Fig. 2 by stacking three layers along the z axis. In this way water molecules inside the four central complexes (see Fig. 2) also experience the field of twelve surrounding complexes, as shown in the x-y projection; *in addition* this sample of 16 complexes is duplicated *twice* by translations both along the +z and along the −z directions, yielding a three-layer sample of 16x3 complexes. As a consequence our final system contains a total of 8640 atoms (not considering the water molecules discussed below); periodic boundary conditions are used.

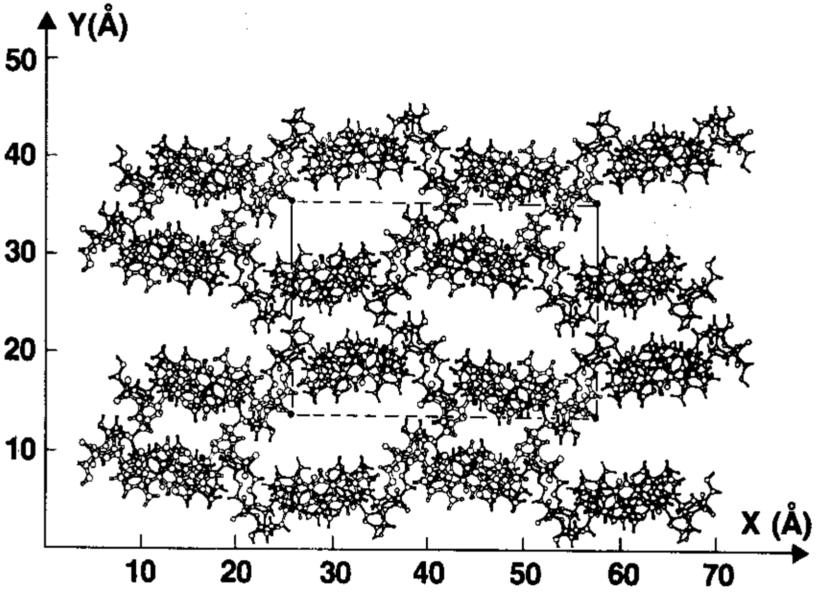

Figure 2. Projection of 16 asymmetric units into the x-y plane. The M.C. simulation considers three such sets, sandwiched in three layers.

In the X-ray diffraction study twenty-five oxygen atoms from water molecules have been reported in each asymmetric unit cell with a maximum resolution of 0.83A. Since our total sample is composed of 48 asymmetric unit cells, we have explicitly considered at least 48x25=1200 water molecules in the Monte Carlo simulations. As explained in detail below, since the exact number of water molecules per unit cell is not known from the X-ray experiment,[2,3] and since the assignment of 50 water molecules per half unit cell is likely underestimated,[8] we have carried out additional simulations with 57, 61, 66 and 77 water molecules per half unit cell; the five computer experiments bracket the number of water molecules per half unit cell. Work is in progress for additional simulations with 50 to 87 water molecules per half unit cell.

We note that the space group determination for the water molecules, experimentally proposed, might be somewhat in question since based mainly on considerations related to the dCpG:PF complexes rather than the dCpG:PF complexes *and* the hydrating water molecules. Let us recall that, strictly speaking, the water molecules can lower the overall symmetry of the solute-solvent system. On the other hand, the problem concerning the definition of the correct space group is likely somewhat premature, because of the still unanswered question on the exact number of water molecules in the unit cell, which leaves the system under study somewhat undetermined. As discussed below we have kept an open mind on the space group assignment. The dCpG:PF complexes have been placed as proposed from the X-ray study.[2] For the water molecules, however, we have assumed a lower symmetry in the unit cell: the Monte Carlo procedure *should be capable of restoring the higher symmetry, if the corresponding conformation is energetically favorable.* Anticipating the analysis which shall follow, the lowering of symmetry consequently means that the very characteristic heptagon of water molecules reported by Neidle, et al.,[2] *do not have to stay symmetric.* The same holds for the water patterns in the major groove. Stated differently, we do not wish to bias our comparison with the crystallographic data by forcing with input constrains the output of our simulations. We recall that the X-ray diffraction data were collected on two different crystals of dCpG:PF complexes which originally yielded two sets of results[3] differing both in the number of water molecules and in the point group assignment. The more highly hydrated crystal appeared less ordered both with respect to the conformation of the molecules and to the structure of the water filaments.[3]

A final note concerns the system's temperature for the experiments. In our computer experiment we have worked both at 273 and 300 K; in this way we should have bracketed the temperature of the X-ray diffraction experiments (which was not reported). In this work we report on the simulations performed at 273 K; we note that the hydration patterns computed at the two temperatures are rather similar.

In conclusion, we hope that the above punctualizations are sufficient to alert against somewhat naive expectations concerning the feasibility of a simple and clear-cut comparison between the results obtained from our *computer experiment* and *from X-ray crystallography.* Anticipating some of the conclusions from our study, we wish to point out that a comparison between the water structures obtained from the two techniques is *not too meaningful* if obtained from models limited to contain as few water molecules as reported in the crystallographic study, being this value too small, as we shall discuss in this work.

Diffraction data from a third crystal and/or re-analyses and refinement of the diffractions using our data could be of interest. Equivalently, our computer experiment can benefit from computational refinements, which are presently in progress. We recall that a strong motivation for this work is due to our opinion that *X-ray diffraction studies* and the type of *computer experiments* below reported *should be performed at the same time, possibly in the same laboratory, in a very complementary mode of operation.*

II. Atom-Atom Potentials

The interactions we shall need are those between a water molecule with a second water molecule, sugar, phosphate, the G and C bases and the proflavine cation, PF$^+$, all of which have been previously reported,[10,11,12,13] with the exception of the latter. We stress that *from our library of pair potentials one cannot obtain with reliability the interaction between a water molecule and PF$^+$*, since the electronic distribution of PF$^+$ is sufficiently different from the electronic distribution of the

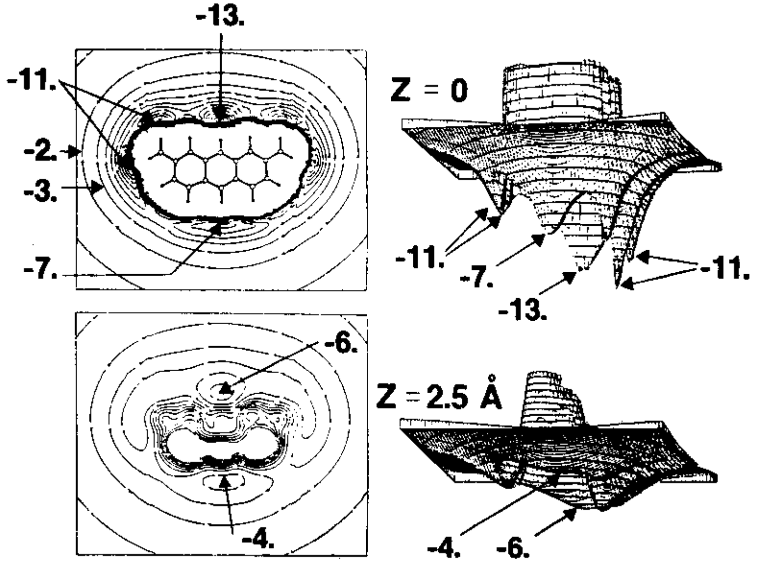

Figure 3. Iso-energy maps for proflavine cation and water: two- and three-dimensional display.

molecules previously analyzed. We have, therefore, performed ab-initio computations on PF^+ interacting with a water molecule, the latter being placed at many positions and orientations relative to PF^+, which is kept rigid; *266 separated points on the interaction energy surface have been computed ab-initio* in the Molecular-Orbitals Self-Consistent-Field approximation. The computations have been performed with our (7s,3p) basis set for the C,N,O atoms and a (4s) set for the H atom;[14] the basis set superposition error has been accounted for by the use of the counter-poised technique.[15] The computer program we have used has been previously discussed.[16]

The 266 ab-initio interaction energies have been fitted to our standard 12-6-1 potential;[4,5] the fitting constants are available upon request.[17] In Figure 3 we present two iso-energy maps for the interaction energy of the system H_2O-PF^+. The water's oxygen atoms have been constrained either in the molecular plane of PF^+ ($z=0$) or in a plane parallel to the molecular plane of PF^+, but shifted by 2.5Å ($z=2.5$). As usual in the iso-energy maps the hydrogen atom's orientations for the water are optimized.[4,5] One can notice for $z=0$ the minima at the hydrogen atoms of the two NH_2 groups (about -10 Kcal/mole), at the central $(N$-$H)^+$ group (about -14 Kcal/mole) and at its opposite site, the C-H site (about -7 Kcal/mole). The contour to contour interval is 1 Kcal/mole.

The fitting of the ab-initio interaction energies, needed to obtain the above analytical atom-atom pair potentials, was carried out to a mean standard deviation of 0.7 Kcal/mole. The overall quality of the fit is shown by the standard plot (fitted/ab-initio) of Fig. 4.

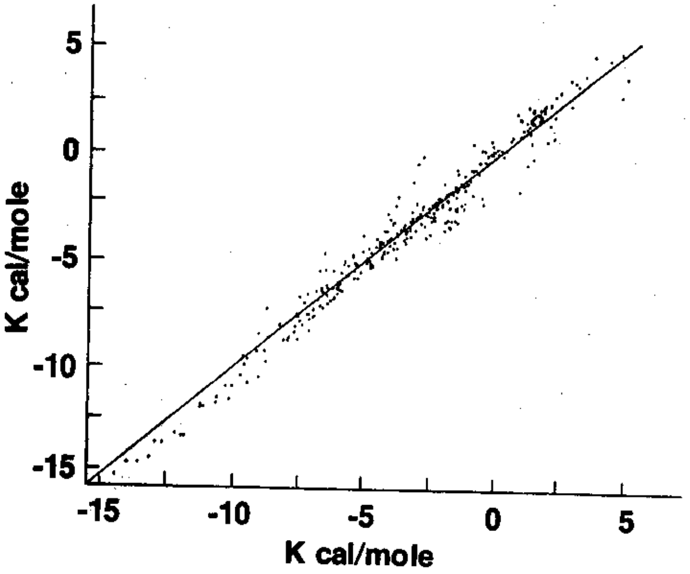

Figure 4. Interaction energies between proflavine and water: fitted energies vs. ab-initio energies (Kcal/mole).

III. Monte Carlo Simulations

In this section we present the results obtained from the Monte Carlo statistical study. We shall not discuss the Metropolis et al.[18] technique, standardly used in statistical Monte Carlo simulations for liquid and solutions, since this is available from many sources.[4,5] We previously have discussed the volume of our sample

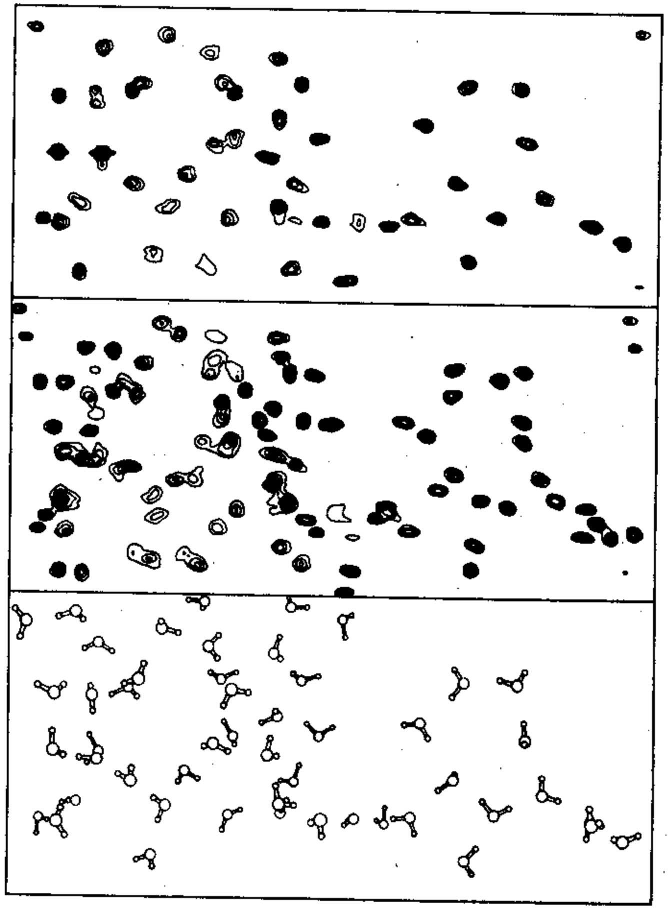

Figure 5. Probability distribution projected into the x-y plane for the oxygen (top), hydrogen (middle) atoms and "averaged" water position for the simulation with 50 water molecules.

(periodic boundary conditions are imposed, as standard). The experiments reported in detail in this work are for 50, 57 or 61 water molecules per half unit cell (we shall hereafter designate as NW the number of simulated water molecules). Experiments for NW=66 and NW=77 are also reported, but discussed only summarily. The results are collected in a number of figures presented in this section. Additional data are available to the interested reader.[17]

Figs. 5-7 display iso-probability distribution maps projected into the x-z plane. These maps (obtained from the Monte Carlo history) display the iso-probability

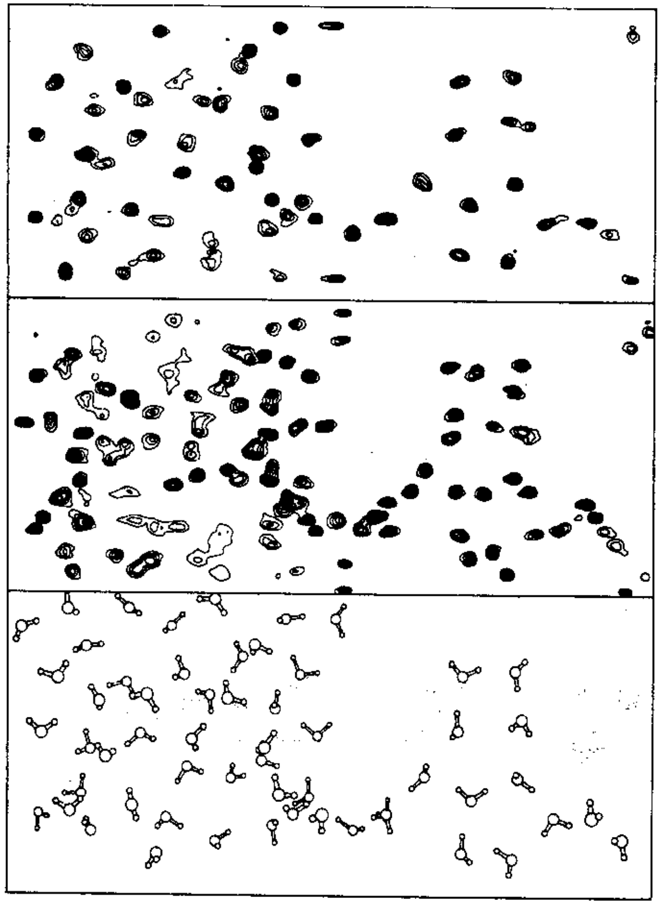

Figure 6. Probability distribution projected into the x-y plane for the oxygen (top), hydrogen (middle) atoms and "averaged" water position for the simulation with 57 water molecules.

contours needed to locate the *oxygen atoms* (top insert) or the *hydrogen atoms* (middle insert) for the water molecules. To help visualization of the data we have added a third insert where the *average positions* (namely, the average values of x, y and z) for the oxygen and hydrogen atoms have been obtained from 150,000 to 200,000 Monte Carlo steps; these average positions are then used, as input, to obtain standard-looking Ortep projections (bottom insert). Thus, from Figs. 5-7 we learn about the diffused or localized nature of the probability distribution; hence, we gain information about the mobility of the water's oxygen and hydrogen atoms. In addition, from the bottom insert in each figure we learn about average positions

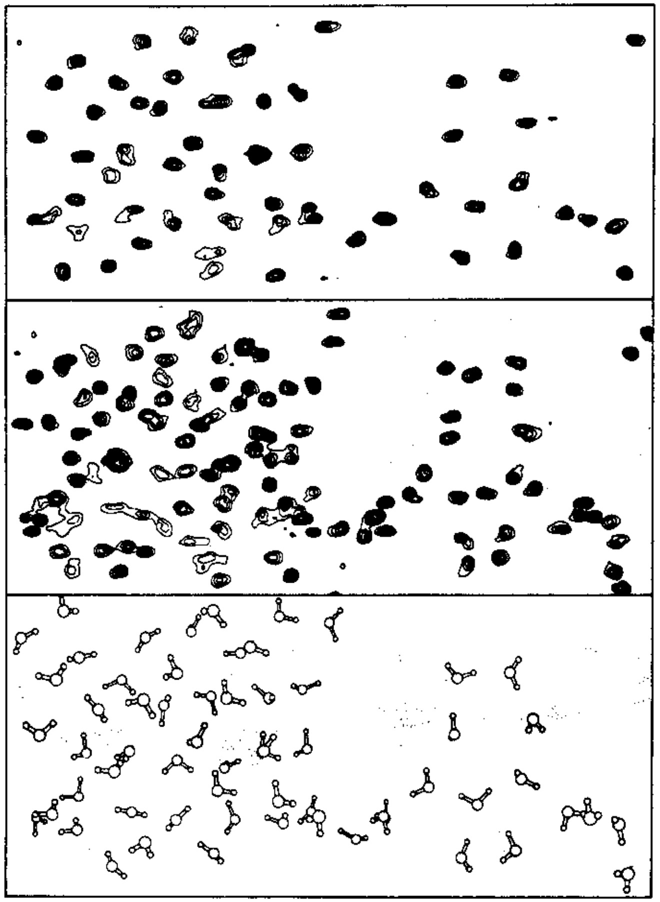

Figure 7. Probability distribution projected into the x-y plane for the oxygen (top), hydrogen (middle) atoms and "averaged" water position for the simulation with 61 water molecules.

of the O-H bond and, therefore about the *average hydrogen-bonding pattern. We recall that, strictly speaking, one cannot discuss the hydrogen-bond between two water's oxygen atoms unless one has evidence on the presence of one hydrogen atom located approximately in between the two oxygen atoms.* We recall, in addition, that for water-water hydrogen-bonds the oxygen-oxygen distance should be between 2.6 and 3.0A. For additional discussions see, for example, Refs. 4 and 5. In Figs. 8-10 we present the oxygen, the hydrogen probabilities and the "average" water molecules in Ortep-type projections, *this time for the x-y plane.*

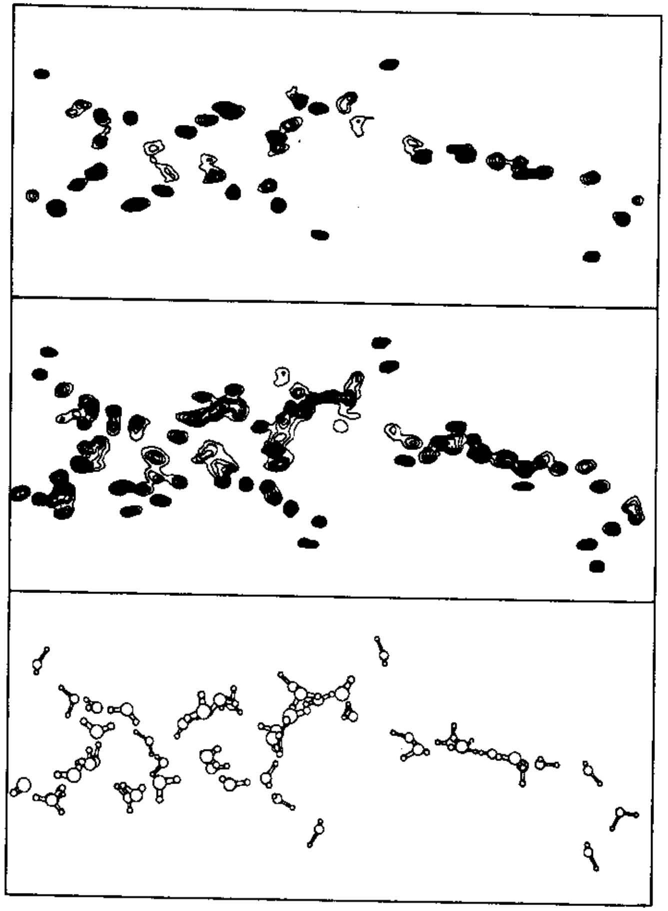

Figure 8. Same as in Fig. 5, but projected into the x-z plane.

In Fig. 11 we report the positions for hydrogen atoms obtained from a simulation where all the *heavy* atoms have been placed at the X-ray coordinates. In this simulation the water's *oxygen atoms are not free to move but the hydrogen atoms can assume any position* (the simulated temperature is 273 K) obtained from rotations. The upper-left insert reports the X-ray oxygen position and the assumed hydrogen-bond pattern; the upper-right insert gives the Monte Carlo probability distributions for the hydrogen atoms; the bottom-right insert reports the *average* orientation for the water molecules and the bottom-left insert reports the Monte

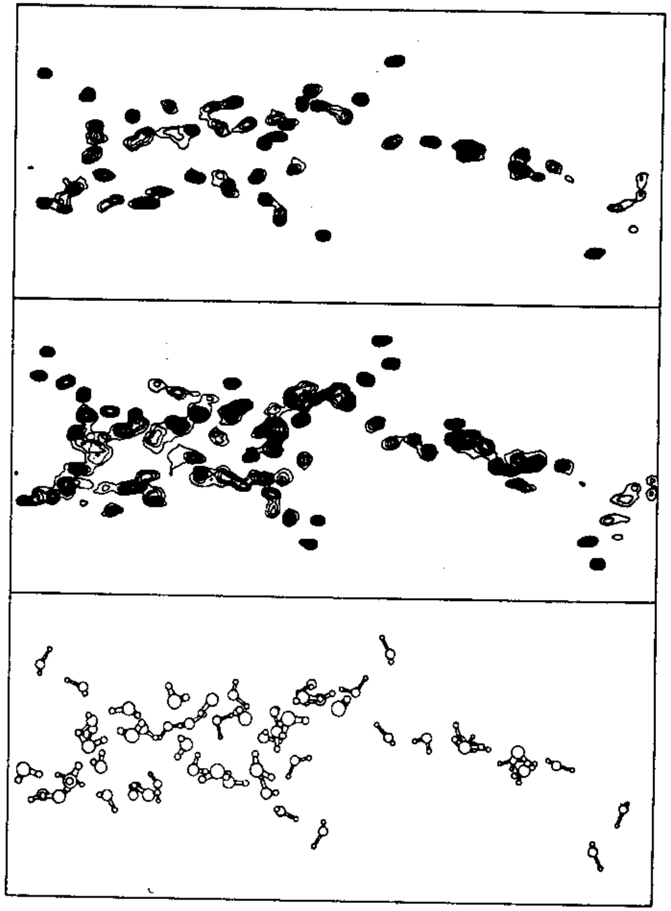

Figure 9. Same as in Fig. 6, but projected into the x-z plane.

Carlo hydrogen bond pattern. The water molecules into the minor groove are the ten given at the left (in each inset) and are characterized by the heptagon-shaped water-water filament. The water molecules to the right appear as if condensed at the borders of the major groove, *with a large empty space in between.* The existence of this large empty space is *very critical* for the water structure determination in the crystal (see below).

In Fig. 12 we reported the values of ΔR, namely the difference in the oxygen atom position predictions obtained either from the X-ray or from Monte Carlo for NW=50, 57 and 61 described in Figs. 5-10. Notice that in the Monte Carlo simulation we have not made use of the *assumed* high space group symmetry, $P2_12_12_1$ adopted in the X-ray study;[2,3] the use of this space group symmetry would bring about a decrease in ΔR biasing, however, the comparison between the crystallographic and the Monte Carlo water-water networks. Therefore, the data of this figure can be taken only as a rough test on the present agreement between the X-ray and the computer experiments. Later, in this work, we shall further discuss the data of this figure.

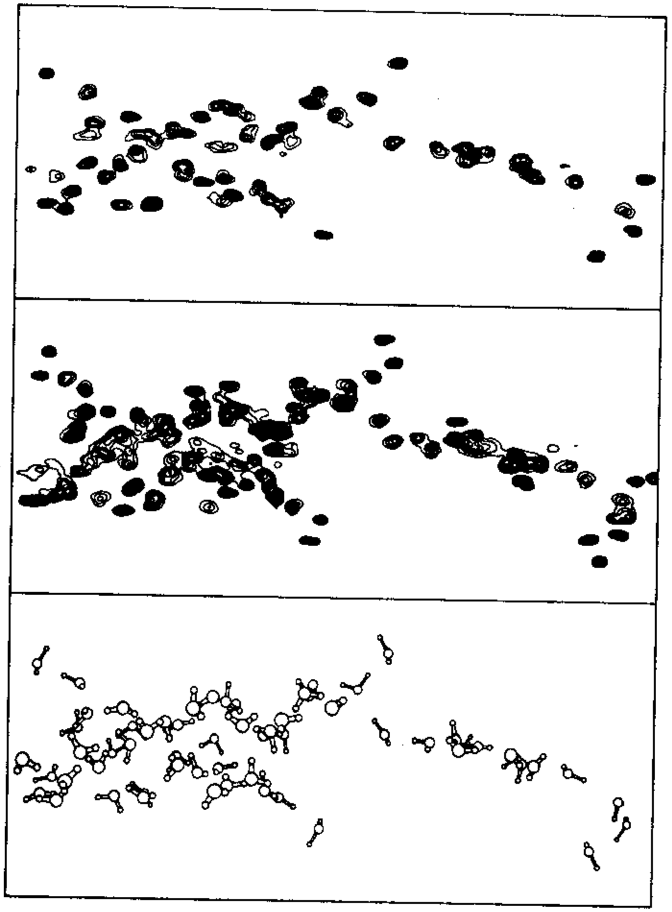

Figure 10. Same as in Fig. 7, but projected into the x-z plane.

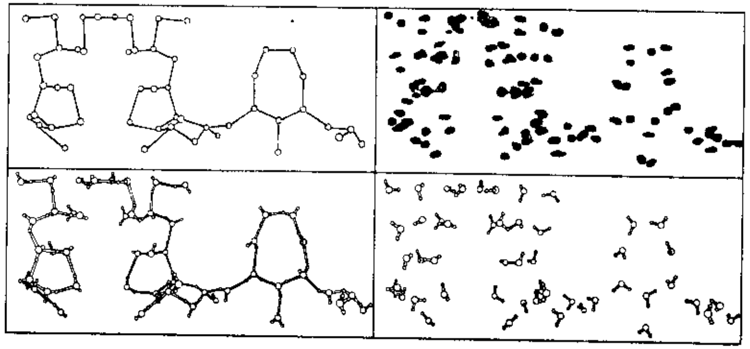

Figure 11. Hydrogen-bonded patterns from the X-ray study (top-left) and from a Monte Carlo simulation (bottom-left). To the right, hydrogen atoms probability distributions (top) and corresponding *average* orientations for water molecules (bottom).

In Figs. 13 and 14 we report the *average* positions of the oxygen atoms for NW=50, 57 and 61 obtained in the Monte Carlo and in the X-ray experiments; the positions are projected both on the x-z plane (Fig. 13) and on the x-y plane (Fig. 14).

Fig. 15 reports the Monte Carlo *average* positions for the oxygen atoms obtained from a simulation with NW=77, side by side with the X-ray positions. The projection is into the x-y plane. The volume of the cell has been sliced into four subvolumes by segmenting the z-axis at z=2.0Å, 5.5Å, 8.5Å and 12.5Å. The two base-pairs are into two 3Å thick slices with z from 5.5Å to 8.5Å and from 12.5Å to 15.5Å; the remaining two slices with z from 2.0Å to 5.5Å and from 8.5Å to 12.5Å, contain the proflavine cations.

Figs. 16-18 report the position and orientation of the 50, 57 or 61 water molecules and either dCpG or proflavine. Fig. 16 reports projections into the x-y plane of 50 water molecules with the oxygen atoms constrained to the positions obtained from X-ray data (top-left), or unconstrained and obtained by a Monte Carlo simulation (top-right); the two insets at the bottom report the Monte Carlo water molecules for NW=57 and NW=61. In the two figures we have reported the G-C base-pairs and part of the PO_4^- groups. The sugars and the proflavines are not reported in order not to further complicate the drawing. Figure 17 reports the same projections as Fig. 16 but with the proflavine cations rather than the base-pairs. In Fig. 18 we

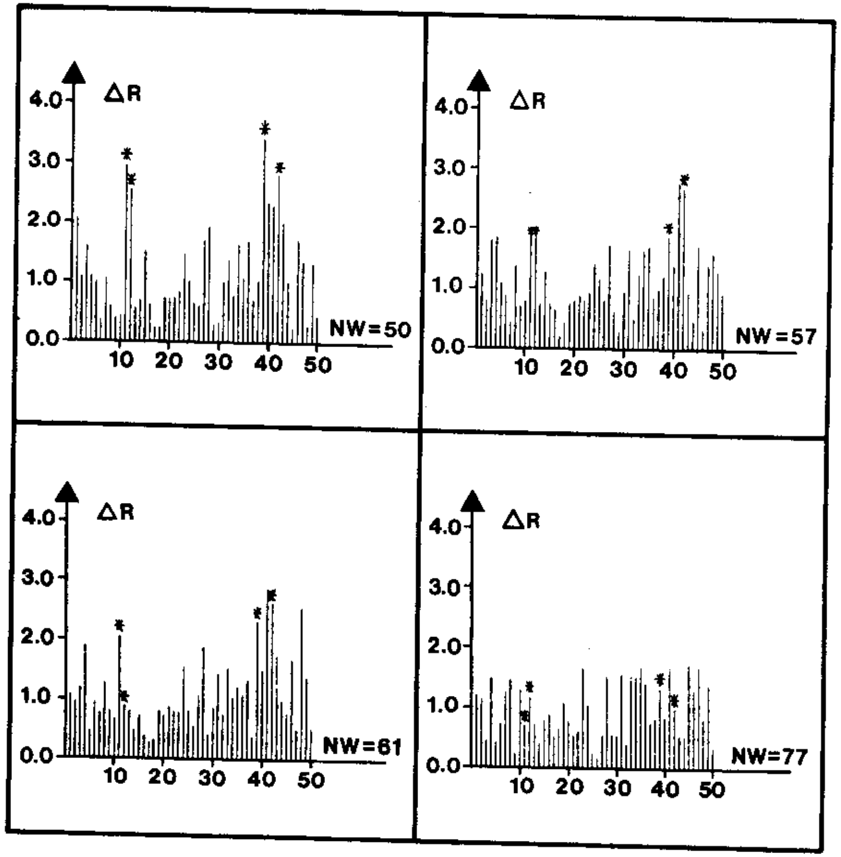

Figure 12. Distance between experimental and computed oxygen atoms.

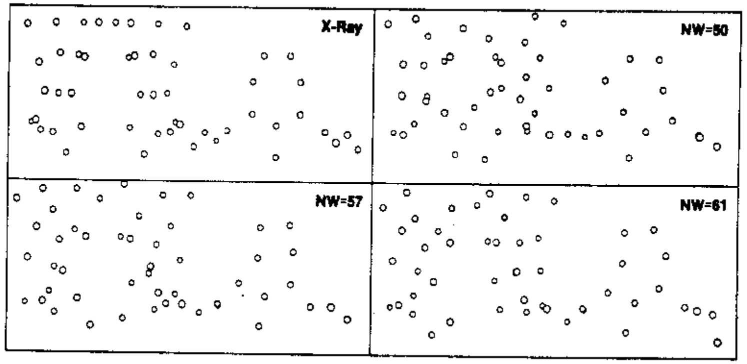

Figure 13. Comparison of the average oxygen positions from X-ray and from Monte Carlo for NW=50, 57 and 61. The projections are into the x-z plane.

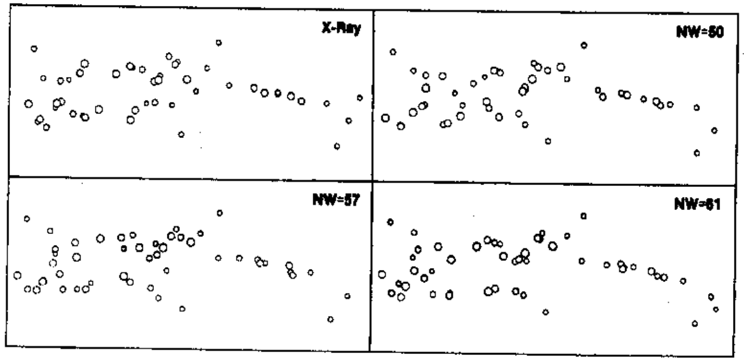

Figure 14. Same as Fig. 13 but for the x-y plane.

project into the x-z plane the 50 water molecules with the oxygen atoms constrained to the X-ray coordinates (left) and the 50 water molecules obtained from the Monte Carlo experiment with free oxygens (right). In the top insets we project the G-C base pairs, in the bottom insets the PF cations. Notice that the minor groove is characterized by the hydrophobic interactions with the C-H groups of PF and the water-attractive NH_2 groups of guanine and the C-O group of cytosine. The major groove is characterized by water-attractive groups: NH_2 of cytosine, N7 and CO of guanine, NH_2 and $(NH)^+$ of PF^+. Therefore, the major groove is more water-attractive than the minor groove. Notice, in addition, that the NH_2 group of PF^+ is connected to the nearby phosphate group by water molecules bridging both groups. The explicit determination of the hydrogen atoms position allows to establish when a water molecule is hydrogen-bonded to another water molecule or to dCpG or to PF^+. From this figure is also evident—for the major groove—not only the large empty hole but also the X-ray very limited hydration of cytosine and of one NH_2 group in proflavine. This assignment, however, can be performed more rigorously by also considering the interaction energies *at the same time.*

Therefore, in Fig. 19 we present the distribution of the interaction energy for each water molecule as obtained from the Monte Carlo history. We recall that the total interaction energy for each water molecule is obtained by summing 1) the interaction energy of the water molecule with all the remaining water molecules and 2) the interaction energy of the same water molecule with all the dCpG:PF complexes.

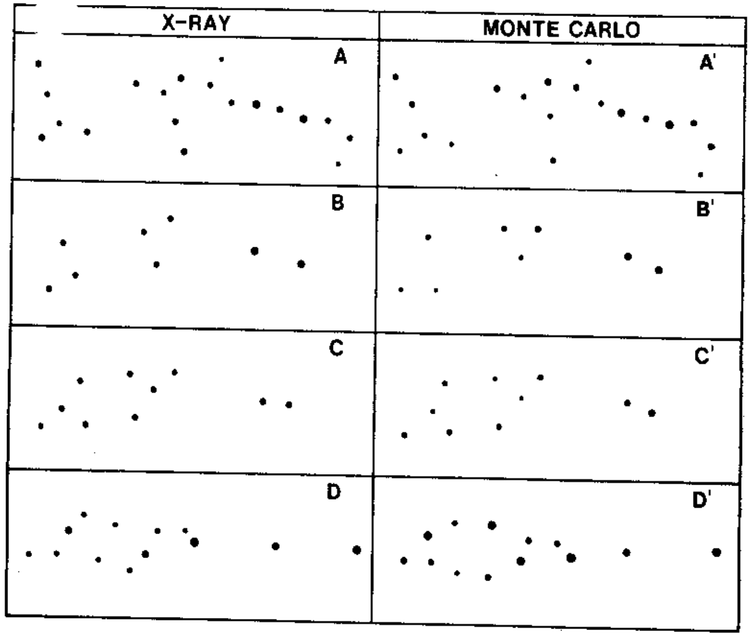

Figure 15. Average position for the oxygen atoms projected into the x-y plane from a Monte Carlo simulation with NW=77 and from X-ray data. The limiting values for Z are 2.0Å to 5.5Å, 5.5Å to 8.5Å, 8.5Å to 12.5Å, and 12.5Å to 15.5Å for A and A', B and B', C and C' and D and D', respectively.

Simply by inspecting the iso-energy maps reported in Fig. 19, one can determine where a given water molecule interacts more with the solute than with the solvent and where a molecule is localized or mobile. Fig. 19 has been obtained for NW=50; equivalent figures have been obtained for NW=57 and NW=61. These are not reported to save printing space, but are made available to the interested reader.[17]

IV. Discussion of the Data and Conclusions

In the following we shall *start* a discussion of the data by accepting as "exact" all the X-ray diffraction conclusions directly based on measurements, namely all the position determinations for the heavy atoms. Subsequently we shall gradually depart from some of the X-ray assumptions, reaching tentative conclusions which might differ from our initial starting point. In this paper *we initiate a hypothetical dialogue which we hope will become "standard" as soon as both X-ray and computer experiments are routinely performed at the same location and at the same time.*

If we start with the assumptions that all the atoms coordinates given in Ref. 2 and 3 are correct and that all the heavy atoms in the crystal are determined, then we can use the Monte Carlo method to determine the *positions of the hydrogen atoms*, for example for the water molecules for which there are "insufficient laboratory evidences."

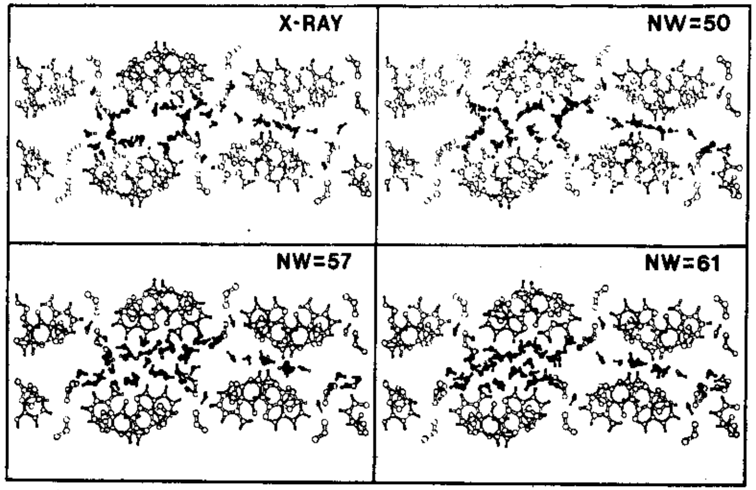

Figure 16. Projections in the x-y plane of the water molecules with the 50 oxygen atoms constrained to the X-ray positions (top-left), and with unconstrained NW=50 (top-right), NW=57 (bottom-left) and NW=61 (bottom-right). Only the G-C bases and part of the PO_4^- groups are reported.

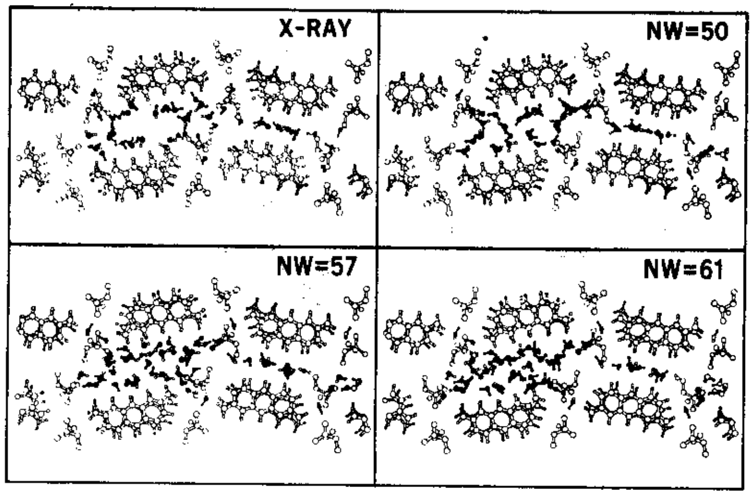

Figure 17. Same as in Fig. 15, this time projecting the PF cations rather than the G-C base-pairs.

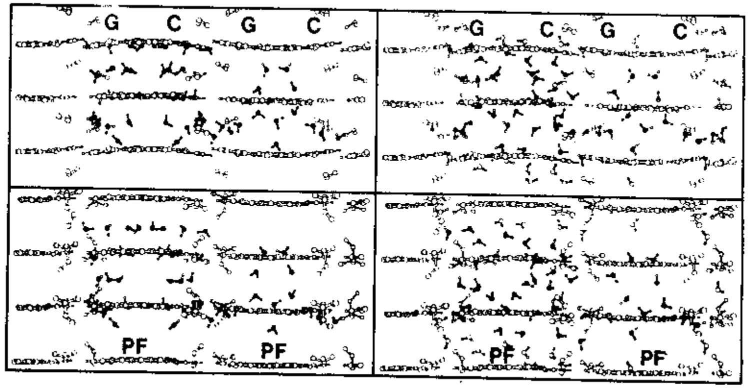

Figure 18. Projections in the x-z plane of 50 water molecules with the oxygen constrained (left) or unconstrained (right) projecting the base-pairs (top) or PF (bottom).

Fig. 11 compares the hydrogen-bond patterns obtained from Refs. 2 and 3 with the one obtained from a Monte Carlo simulation where i) all the heavy atoms are placed as from Refs. 2 and 3, ii) hydrogen atoms are added to the dCpG:PF complexes and to the oxygen atoms of the water molecules and iii) the orientation (but not the oxygen atom position) for the water molecules are varied randomly many times and analyzed in the standard Monte-Carlo Metropolis way (the simulated temperature is 273K). Notice in addition that the water molecule located at the bottom of the heptagon has been assigned to be in two possible sites (see Ref. 3) each one with 50% of fractional occupation. In our simulation we have placed the oxygen atom at a site located midway between the two partially occupied ones, since the reported oxygen-oxygen distance between the two fractional waters is only 1Å.

In the Monte Carlo simulation, the $P2_12_12$ space group should not be imposed for the water molecules not only because two waters in the minor groove are located along a symmetry axis, but also because some waters are near to symmetry axes.

The top-left inset reports the X-ray "hydrogen-bond pattern;" the top-right inset reports the Monte Carlo probability distribution for the hydrogen atoms. The bottom-left inset reports the hydrogen-bond pattern (obtained from the computed *average positions of each water molecule*, bottom-right insets) with an explicit visualization for the water-water hydrogen bonds.

The resulting Monte Carlo pattern is essentially equal to the one generated from the X-ray data as clearly visible by comparing the two insets on the right side of the figure. Note that *superficially* the appearance might seem somewhat different, since the hydrogen atoms often make *"bent"* hydrogen bonds. Notice also that the probability distribution of the hydrogen atoms are rather localized, indicating low rotational freedom, possibly an artifact of our constraints imposed on the oxygen positions. Concerning the hydrogen-bond pattern, it is gratifying to find from experiments more and more evidence of the water-water filaments, which we have been predicting for some time.[5,6,7]

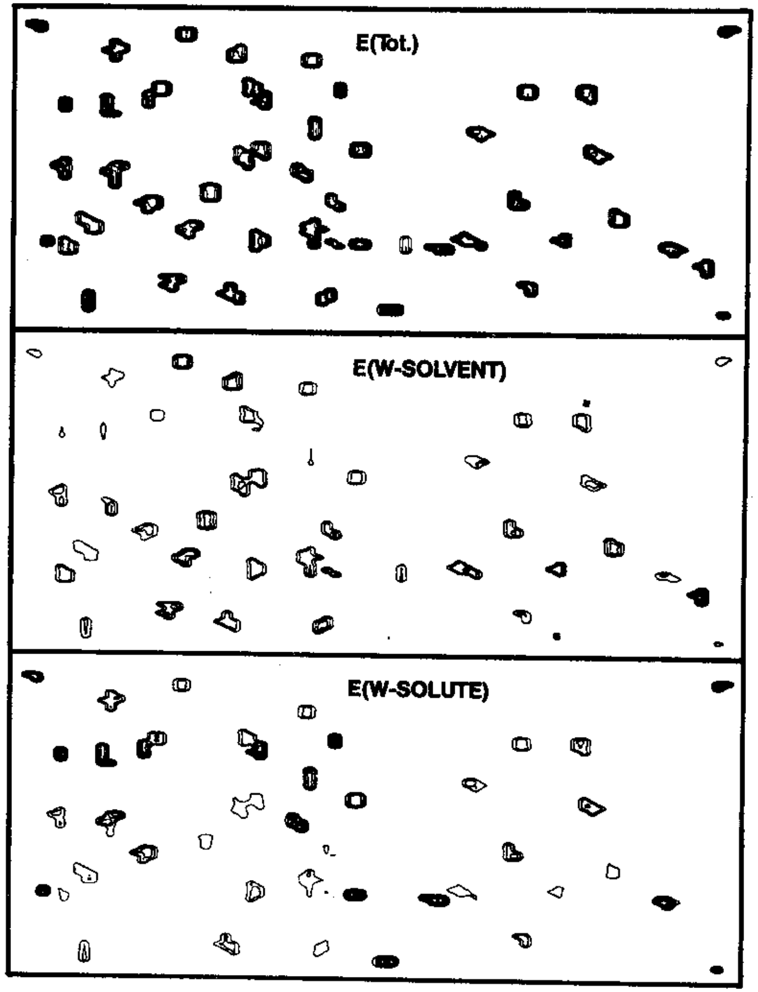

Figure 19. Iso-energy maps for the water molecules from the Monte Carlo experiment with NW=50. Top: total interaction energy for each molecule of water. Middle: interaction of one water with all the other water molecules. Bottom: interaction of each water with the dCpG:PF complexes.

There is, however, a noteworthy difference: in the Monte Carlo experiment we have assumed a lower space group symmetry, $P2_1$, than in the X-ray analyses and this brings about the possibility of finding asymmetrical orientations in water molecules (when energetically more preferable than symmetrical orientations). Typical examples of asymmetry are the orientations of the water molecules which are in the very characteristic heptagon (bottom inset, right side of Fig. 11). The resulting averaged orientations of the water molecules *do not seem to support the* $P2_12_12$ *space point symmetry,*[2,3] *for the hydrated system.* This conclusion is taken as a strong indication that in the crystallographic study[3] a number of water molecules were undetected.

The computed *average interaction energy* for one water molecule, W, interacting with the remaining water molecules is E(W-solvent)=−19.9±0.2 Kj/mole; the average interaction energy of a water molecule with the complexes (solute) is E(W-solute)= −39.2±0.2 Kj/mole. The above computed values tell us that most of the water molecules interact strongly with the solute and much less with the solvent. These conclusions, coupled with the observation that the water-water interaction is computed as about one half as found in bulk water, tell us that the X-ray water positions correspond mainly to hydrophilic interactions and that there is no "bulk-type water." The size of the major groove, however, is large enough to support the expectation of more water molecules than those reported from the X-ray experiment. Indeed we have seen in Fig. 11 and from Figs. 16-18 a rather large region of *empty space,* without water molecules in the major groove. Relaxing the positional constraint on the oxygen atoms of the 50 water molecules, a few water molecules might migrate to the empty space vacating the sites determined by the X-ray experiment; notice in addition that the molecules once in the previously empty region are expected to have a rather diffuse probability, namely have a rather high mobility.

To study this possibility more deeply, we now relax a conclusion previously accepted from the X-ray experiment; we allow the water molecules not only to assume random orientations but also random positions. This Monte Carlo simulation has been performed both at 273K and 300K; here we shall report data only from the 273K experiment.

Some of the results from this Monte Carlo experiment are presented in Figs. 5 and 8. In Figs. 13 and 14 we give the computed and the X-ray oxygen atom positions, projecting the atoms either into the x-y or in the x-z planes and in Figs. 16-18 we have presented the projections reporting in addition the base-pairs, proflavine and part of the phosphate groups.

From Figs. 5, 8 and 19 we learn that about 20 water molecules have a *very localized probability distribution,* whereas the remaining ones have a somewhat diffuse distribution. *All the 10 water molecules in the minor groove coincide with the X-ray predictions* (the experimental resolutions is 0.83Å and the average deviation between X-ray and Monte Carlo is given by ΔR=0.70Å). Let us now point out the major differences which mainly concern the water molecules in the major groove.

In the major groove about 12 water molecules are within 0.3 to 0.9Å from the X-ray experiment; 19 are within 0.9Å and 1.9Å and 9 between 1.9Å and 3.4Å. We recall that from the X-ray experiment there is a large hole approximately at the "center" of the major groove; by symmetry this hole is as long as the dimensions of the crystal, since it grows along the z direction. The average ΔR for the fifty water molecules is 1.17Å.

Therefore, nine molecules out of 50 have migrated by about 2 to 3Å away from their "X-ray position," into the "X-ray empty space."

In this experiment the average energetics for the system of water molecules is as follows: E(W-solute)=−52.0±0.1 Kj/mole and E(W-solvent)=−19.2±0.1 Kj/mole. Thus, in average the "Monte Carlo positions" for the 50 water molecules (at a temperature of 273K) are computed to correspond to a conformation *more stable by about 10 Kj mole per molecule of water relative to the conformation* proposed by the X-ray experiment. Note that the value of this relative stabilization is far larger than the mean standard deviation, which is about 0.1 Kj/mole per molecule of water.

One more observation: the lowering of the space group symmetry had a minor consequence on the previous Monte Carlo experiment, since only rotation of the water molecules where allowed. In this experiment the consequences are more dramatic; for example, the beautiful heptagonal structure of water molecules in the minor groove becomes rather irregular. Would have we retained the $P2_12_12$ space group, we would have obtained a more regular structure, biasing, however the simulation by forcing agreement with the crystallographic data.

Tentatively, one would like to conclude from this experiment that there is a rather large channel-like central cavity in the major groove where water molecules could go; more quantitatively one would assume that *at least 9 molecules of water are interested in this migration.* This would lead to the tentative conclusion that NW should be about at least as large as 59, assuming both the X-ray and the Monte Carlo data to be equally valid. The latter assumption does not lead to contradictions: *the X-ray cannot see mobile water molecules, and the Monte Carlo selects as more probable those configurations which are energetically more stable.*

To explore *preliminarily* this supposition we have performed additional Monte Carlo experiments with NW=57, 61 (see Figs. 6, 7 and 9, 10 respectively). Having determined, on the basis of considerations of the interaction energy that the system of water molecules can be even larger, we have extended the expermentation to NW=66, 77 and 87. We are now performing systematic experiments with 50≤NW≤80. These simulations are only preliminary. Notice that we have placed additional molecules of water in the minor groove (the probability for a molecule of water to pass from groove to groove is rather low) in order to study whether more water can be fitted into the small groove even if the agreement with the X-ray data is already satisfactory.

The energetic of the two Monte Carlo experiments with NW=57 and NW=61 are: E(W-solvent)=−20.9 Kj/mole and −21.0 Kj/mole for NW=57 and 61 respectively and E=(W-solute)=−47.8 Kj/mole and −46.4 Kj/mole for NW=57 and 61, respectively. Notice that the total energy for the system is given by the product [E(W-solute)+E(W-solvent)]NW; the NW factor combined with the increase in the average water-water attraction more than compensates for the small loss in the E(W-solute). Thus, on the basis of an energy minimization criterion, one could expect an NW larger than 61. This preliminary conclusion is, however, presently put forward as a tentative one. However, the preliminary data from the Monte Carlo experiments with NW=66 and 77 do corroborate the above preliminary conclusion.

The most convincing demonstration in favor of a large value for NW comes from a simulation with NW=77: the computed average positions for the oxygen atoms yield patterns *in notable agreement with the crystallographic data as clearly shown in Fig. 15.* The value of ΔR, the average difference between the oxygen atom positions from X-ray and Monte Carlo, is about 0.92Å with a Monte Carlo resolution of about 0.33Å. For the water molecules in the minor groove ΔR is 0.52Å; we recall that the X-ray resolution is 0.83Å. We refer to Fig. 12 for a display of ΔR for NW=50, 57, 61 and 77.

The trends in the computed ΔR, in the total interaction energy, in the average water-water contact distances point out that $65 \leq NW \leq 75$. This conclusion is reached from analysis on preliminary data for simulations with NW=82 and NW=87. The above trend points out that one should converge to an average value of ΔR not dissimilar from the one above reported for the molecules of water in the minor groove.

We are presently refining our computer experiments to firmly establish the value of NW and to determine the effects, even if small, of the variations of temperature from 273 K to 300 K. All this will take some time, since these Monte Carlo simulations and the concomitant analyses represent a rather heavy computational task.

However, we feel to have essentially reached our main objective namely we have shown with the type of computer experiments we have since long advocated[4-7] one can obtain *relevant and complementary conclusions* to laboratory experiments, for example crystallographic studies dealing with complex chemical structures of interest to current biophysical problems. The wealth of *quantitative structural, dynamical* and *energetic-type details* which one can routinely obtain has been demonstrated by reporting the large number of figures. Finally we would like to recall the full generality of the computational experiment.

In this work we have used an IBM-4341/FPS-164 coupled system, using computer programs developed in our department.

References and Footnotes

1. Westhof, E. and Sundaralingam, M., *Proc. Natl. Acad. Sci. USA 77*, 1852 (1980) and Swaminathan, P., Westhof, E. and Sundaralingam, M., *Acta Cryst. B 38*, 515 (1982)
2. Shieh, H.S, Berman, H.M., Dabrow, M. and Neidle, S., *Nucleic Acids Res., 8*, 85(1980)
3. Neidle, S., Berman, H. M. and Shieh, H.S., *Nature 288*, 129 (1981)
4. See for example: Clementi, E., *Determination of Liquid Water Structure, Coordination Numbers for Ions and Solvation for Biological Molecules, Lecture Notes in Chemistry,* Vol. 2, Springer-Verlag, Berlin, Heidelberg, New York (1980)
5. Clementi, E., *Computational Aspects for Large Chemical Systems, Lecture Notes in Chemistry,* Volume 19, Springer-Verlag, Berlin, Heidelberg, New York (1980)
6. Clementi, E. and Corongiu, G., *Int. J. Quantum Chem. 22*, 595 (1982); Clementi E., and Corongiu, G., *Biopolymers 21*, 793 (1982)
7. Clementi, E. in *Structure and Dynamics: Nucleic Acids and Proteins* Clementi, E. and Sarma, R.H., eds., Adenine Press, pp. 321-364 (1983)
8. Private communication (Conversation in Biomolecular Stereodynamics, Albany, NY, June 7-10, 1983) from Neidle, S. and Berman, H. M. to two of us (K.K. and G.C.)
9. See for example Lie, G.C., Clementi, E. and Yoshimine, M., *J. Chem. Phys. 64*, 2314 (1976); Clementi, E., Kistenmacher, H., Koloz, W. and Romano, S., *Theoret. Chim. Acta (Berl.) 55*, 257 (1980) and Corongiu, G. and Clementi, E., *Int. J. Quantum Chemistry,* (Sanibel Simposia Volumes), in press.
10. Matsuoka, O., Clementi, E. and Yoshimine, M., *J. Chem. Phys. 64*, 1351 (1976)
11. Corongiu, G. and Clementi, E., *Gazz. Chim. It. 108*, (5-6), 273 (1978): note that the charges on the sugar are those recently reported by Corongiu, G. and Clementi, E., *Int. J. Quantum Chem. 9*, 213 (1982)
12. Clementi, E., Corongiu, G. and Lelj, F., *J. Chem. Phys. 70*, 3726 (1979): note that the charges on the phosphate are those recently reported by Corongiu, G. and Clementi, E., *Int. J. Quantum Chem. 9*, 213 (1982)
13. Scordamaglia, R., Cavallone, F. and Clementi, E., *J. Am. Chem. Soc. 99*, 5545 (1977): note that the charges on the bases are those recently reported by Corongiu, G. and Clementi, E., *Int. J. Quantum Chem. 9*, 213 (1982)
14. Gianolio, L. and Clementi, E., *Gazz. Chim. It. 110*, 179 (1980)
15. Boys, F.S. and Bernardi, F., *Mol. Phys. 19*, 553 (1970)
16. Ortoleva, E., Castiglione, G. and Clementi, E., *Comp. Phys. Comm. 19*, 337 (1980) and Habitz, P. and Clementi, E., *Comp. Phys. Comm. 29*, 301 (1983)
17. The interested reader should request the IBM Research Report, POK-26, June, 1983, from one of us (G. C.)
18. Metropolis, N., Rosenbluth, A. W., Teller, A. H. and Teller, E., *J. Chem. Phys. 21*, 1078 (1953)

Date received: August 2, 1983 on computer tape as EBCDIC images