Water and Ion Binding Around RNA and DNA (C,G) Oligomers

Pascal Auffinger and Eric Westhof*

Institut de Biologie Moléculaire et Cellulaire du CNRS
Modélisations et Simulations des Acides Nucleiques, UPR 9002, 15 rue René Descartes
67084, Strasbourg Cedex
France

The dynamics, hydration, and ion-binding features of two duplexes, the A(r(CG)12) and the B(d(CG)12), in a neutralizing aqueous environment with 0.25 M added KCl have been investigated by molecular dynamics (MD) simulations. The regular repeats of the same C=G base-pair motif have been exploited as a statistical alternative to long MD simulations in order to extend the sampling of the conformational space. The trajectories demonstrate the larger flexibility of DNA compared to RNA helices. This flexibility results in less well defined hydration patterns around the DNA than around the RNA backbone atoms. Yet, 22 hydration sites are clearly characterized around both nucleic acid structures. With additional results from MD simulations, the following hydration scale for C=G pairs can be deduced: A-DNA < RNA (+3 H2O) and B-DNA < RNA (+2 H2O). The calculated residence times of water molecules in the first hydration shell of the helices range from 0.5 to 1 ns, in good agreement with available experimental data. Such water molecules are essentially found in the vicinity of the phosphate groups and in the DNA minor groove. The calculated number of ions that break into the first hydration shell of the nucleic acids is close to 0.5 per base-pair for both RNA and DNA. These ions form contacts essentially with the oxygen atoms of the phosphate groups and with the guanine N7 and O6 atoms; they display residence times in the deep/major groove approaching 500 ps. Further, a significant sequence-dependent effect on ion binding has been noted. Despite slight structural differences, K+ binds essentially to GpC and not to CpG steps. These results may be of importance for understanding various sequence-dependent binding affinities. Additionally, the data help to rationalize the experimentally observed differences in gel electrophoretic mobility between RNA and DNA as due to the difference in hydration (two water molecules in favor of RNA) rather than to strong ion-binding features, which are largely similar for both nucleic acid structures.

*Corresponding author

Keywords: molecular dynamics; RNA; DNA; hydration shell; gel electrophoresis

Introduction

The aqueous solvent plays a central role in the maintenance of the three-dimensional architecture of biopolymers. It has been shown by experimental and theoretical methods that water molecules, even located at long distances from the solute (>10 Å), are structurally important (Auffinger et al., 1996; Israelachvili & Wennerström, 1996). These apparently unstructured hydration shells are involved in the formation of hydration forces at play in folding and recognition events and should be studied thoroughly (Sorenson et al., 1999). On the other side, a precise knowledge of the shape of the primary hydration shell at the nucleotide and base-pair level is required. Indeed, ligands (nucleotides, amino acids, drugs, . . .) interact with nucleic acids either by direct contacts, thereby replacing some water molecules from their locations in the first hydration shell, or by forming water-mediated interactions. Most commonly, ligands establish both types of contacts with nucleic acids.

Crystallography is the method of choice for studying the primary hydration shell of nucleic acids and several reviews have surveyed the field.
The shapes of the hydration shells around DNA bases (Schneider & Berman, 1995). DNA phosphate groups (Schneider et al., 1998) and RNA C=G, A=U, G=U base-pairs (Auffinger & Westhof, 1998c) have been investigated on the basis of statistical analysis of high-resolution crystallographic structures. Yet, the main characteristics of the hydration shells obtained by crystallographic methods are not always completely determined, and X-ray techniques provide only limited information on the dynamics of the bound water molecules. In addition, the hydrogen bond patterns formed between the solvent and the solute molecules have often to be guessed, since hydrogen atoms are not observable. Molecular dynamical (MD) simulations allow the completion of the structural and dynamical information gained by experimental methods. Several recent MD studies have been devoted to the characterization of the properties of the hydration shells around RNA and DNA structures (Auffinger & Westhof, 1997a, b; Cheatham & Kollman, 1997; Duan et al., 1997; Young et al., 1997b; Cheatham et al., 1998; Feig & Pettitt, 1999a). Hydration patterns were often described at the level of the global nucleic acid structure and many heterogeneous sequences have been investigated with the principal aim of characterizing the minor groove “spine of hydration” first observed for the d(CGCGAATTCGCG)2 dodecamer (Kopka et al., 1983). Yet, the understanding of the physicochemical basis for changes in hydration induced by the addition or removal of a single 2'-OH group and by the U → T modification in the structural and functional properties of nucleic acids is of great importance (Wang & Kool, 1995; Norberg & Nilsson, 1996; Cheatham & Kollman, 1997). Such an understanding should provide useful hints to recognition and folding problems as well as explanations to the fact that, for a similar sequence, RNA duplexes are thermodynamically more stable than DNA duplexes (Lesnik & Freier, 1995; Conte et al., 1997; Kankia & Marky, 1999).

In an earlier study, we described the preferred orientations and the hydration properties of the 2'-OH group (Auffinger & Westhof, 1997b). The present study focuses on the characterization of the similarities and differences in the hydration and ion-binding properties of regular RNA and DNA duplexes. To this end, two 2.4 ns MD simulations of two regular helices in an aqueous environment with 0.25 M KCl added, derived from the Arnott fiber models (Arnott et al., 1976), the A-form RNA r(CG)12 duplex and the B-form DNA d(CG)12 duplex, were generated (Figure 1). The choice of regular sequences containing C=G pairs may appear trivial compared to “more biologically relevant” structures. Yet, Watson-Crick base-pairs are the single most important secondary structural motif in nucleic acids. Short strands of alternating C=G pairs can be found in numerous RNA and DNA sequences, and the grooves of C=G pairs are involved in a great variety of intermolecular contacts in regular triple helices, base triples, and complexes formed between protein or drugs and nucleic acids. From a computational point of view, the aims of selecting sequences containing a regular repeat of C=G base-pairs are multiple. First, the comparison of such sequences involves only chemical changes associated with the presence or absence of the 2'-OH group and do not consider the modifications associated with the presence of U (RNA) or T (DNA) bases. Second, among nucleic acid duplexes, those containing only C=G base-pairs are thermodynamically the most stable (SantaLucia, 1998; Xia et al., 1998) and, consequently, are supposed to be structurally less flexible than A-U or A-T-containing sequences, a fact that therefore, should limit problems associated with the sampling of the conformational space. Third, for a regular sequence, it is possible to calculate average values for various properties of the C=G base-pairs in a more precise way than by analyzing sequences where C=G pairs are located in a heterogeneous environment. This is crucial for investigating dynamical features that need long equilibration times, such as those related to the ionic atmosphere. The use of regular or semi-regular (Feig & Pettitt, 1998b, 1999a) sequences is similar in spirit to applications of the multiple molecular dynamics (MMD) method that have been designed to extend the sampling of the multiple conformational space of specific structural motifs, a recurrent problem in MD simulations (Auffinger et al., 1995; Auffinger & Westhof, 1997a; Caves et al., 1998; Worth et al., 1998).

Besides the structural aspects related to the nucleic acid hydration shells, the question of whether and where ions may bind to nucleic acids is of great importance, as it is well appreciated that ion coordination sites are potential binding sites for charged groups of proteins or drugs (Draper & Misra, 1998; Hermann & Westhof, 1998). It has long been stated that the primary hydration shell of nucleic acids is impermeable to ions and, indeed, only a limited number of divalent or monovalent ions in direct contact with the nucleic acids have been experimentally observed. MD simulations have challenged this view, since it has been reported that direct contacts between monovalent ions and nucleic acids could occur in RNA and DNA grooves (Mohan et al., 1993; Weerasinghe et al., 1995b; Auffinger & Westhof, 1997a; Cheatham & Kollman, 1997; Young et al., 1997a, b; Lyubartsev & Laaksonen, 1998; Feig & Pettitt, 1999b). These observations have been rationalized and several binding sites have been proposed in both DNA grooves depending on the base-pair sequences (Young et al., 1997a). From the experimental side, recent crystallographic data have confirmed that monovalent ions may partially intrude into the nucleic acid grooves (Shui et al., 1998; Tereshko et al., 1999), while NMR evidence for NH$_3^+$ (Hud et al., 1999) and other monovalent...
ion groove binding (Denisov & Halle, 2000) has been reported. Further, well-characterized K$^+$ (Basu et al., 1998) and Na$^+$ (Su et al., 1999) binding pockets have been described in RNA crystal structures. Thus, despite some discussions concerning the significance of effects related to ion-binding events (Chiu et al., 1999; McFail-Isom et al., 1999; Beveridge & McConnell, 2000), there is little doubt that some hydration sites localized in the first coordination sphere of nucleic acids are partially occupied by monovalent ions. In the present MD study, potassium ions were considered instead of the more common sodium ions. The main reasons for choosing potassium ions are, first, that they represent one of the dominant ionic species in the intracellular medium and, secondly, that it has been demonstrated that potassium ions are required in some RNA molecules for structural and functional integrity (Wang et al., 1993; Gluick et al., 1997; Basu et al., 1998; Draper & Misra, 1998).

Another aim of this study is to provide a structural basis for the observed difference in gel electrophoretic mobility between RNA and DNA duplexes (Ratmeyer et al., 1994; Lesnik & Frejer, 1995; Bonifacio et al., 1997; Hagerman, 1997). Despite the fact that both types of molecules display the same net charge, RNA duplexes migrate more slowly than DNA duplexes under various ionic conditions. This behavior has been attributed to different ion distributions around A-form and B-form duplexes (Bonifacio et al., 1997). However, the fact that RNA molecules bear an additional 2'-OH group compared to DNA molecules and are, therefore, supposed to be more hydrated (Egli et al., 1996) could explain as well the observed differences in mobility. The data on the hydration and ionic atmosphere of the RNA and DNA duplexes provided by the MD simulations should help in unraveling the origins of this behavior.

Figure 1. Canonical A-form structure of the r(CG)$_{12}$ helix (left) and canonical B-form structure of the d(CG)$_{12}$ helix (right) in their simulation box. Middle, sequence of the RNA and DNA duplexes analyzed in this study. Only the central 12 base-pairs (in yellow) were used for the determination of the hydration and ion-coordination sites. Each box is filled with 72 K$^+$ (in blue) and 26 Cl$^-$ (in green). For clarity, the water molecules filling the simulation boxes are not displayed.
Results

Global structural features

The r(CG)$_{12}$ and the d(CG)$_{12}$ alternating oligomers retain their overall structural characteristics during the 2.4 ns trajectories. No single base-pair opening event was observed during the simulations of either oligomer. For the RNA duplex, the RMS deviation profiles and the snapshots extracted from the MD trajectories show that the structures remain close to the initial A-form shape, while they stay constantly at the high RMS value of 6 Å from a hypothetical B-form RNA structure (Figure 2). The superposition, at the base-pair level, of the initial and the average structure results in a low RMS deviation value of 0.5 Å, despite a 1.0 Å increase of the initial 18.0 Å interphosphate distance (Figure 3). The sugar pucker stays in the C3'-endo domain with low angle and amplitude variations. The average value of the pseudorotation phase angle $P$ ($15^\circ$) is close to the canonical C3'-endo value ($18^\circ$).

On the contrary, although no clear central hole characteristic of an A-form DNA can be seen in the

![Figure 2](image-url)

Figure 2. RMS deviations and snapshots extracted from the MD simulations of the r(CG)$_{12}$ and d(CG)$_{12}$ duplexes. Top, RMS deviation profiles from the starting structures. The red lines correspond to the complete 24 base-pair structures and the black lines to the central 12 base-pair segments. Additionally, for the RNA and the DNA simulations, RMS deviations with respect to a B-form and an A-form structure are given, respectively (the RNA B-form structure is ‘hypothetical’). The horizontal line marks the RMS deviations between the starting A and B-form structures. The time during which positional constraints have been applied is emphasized by a red gradation. The end of the equilibration period is marked by a broken line. Bottom, superposition of 20 structures covering 400 ps of MD extracted from the beginning and the end of the RNA (left) and DNA (right) MD simulations. While the A-RNA structures remain essentially of the A-type, the B-DNA moves towards a form intermediate between A and B.
overlay of MD structures (Figure 2), it appears that the final structures are intermediate between an A-form and a B-form DNA, as further indicated by the convergence of the RMS deviation profiles (MacKerell & Banavali, 2000) calculated with respect either to the initial B-form structure or to a standard A-form helix (Figure 2). The calculated RMS deviation at the base-pair level is larger for the DNA structure (0.7 Å) compared to the RNA structure (0.5 Å) and the interphosphate distance increases on the average by 1.3 Å (Figure 3). The DNA sugar puckers remain essentially in the C2'-endo domain with some incursions in the O4'-endo domain and very few switches to the C3'-endo domain. The mean value of the pseudorotation phase angle $P$ for the sugar population centered on the C2'-endo domain is close to 145°, which is less than the canonical value (162°), and is characteristic of the Cornell et al. force-field (Cheatham et al., 1999). The histogram representing the $\tau_m$ values is broader for DNA than for RNA, reflecting the larger mobility of the sugars. The mean value is 2°.
lower (41.5° for RNA compared to 39.5° for DNA). Close to 180 backbone dihedral angle transitions (≈2 transitions residue^{-1} ns^{-1}) involving 35 residues were observed for DNA, while for RNA, only 12 backbone dihedral angle transitions (≈0.2 transitions residue^{-1} ns^{-1}) involving five residues among the 48 of the duplex were noted during the last 1.5 ns of simulation. The interphosphate distance displays a larger range of variability for DNA than for RNA (Figure 3) but, similarly to RNA, is larger by ≈1 Å than the initial interphosphate distance derived from the fiber model structure.

The well-behaved structural characteristics displayed by the two duplexes along the MD trajectories allow for an in-depth analysis of their ion-binding and hydration properties. In the following, we focus our analysis on the central 12 base-pairs of each duplex (Figure 1) in order to avoid the inclusion of possible end-effects in the statistics.

**Ion-binding sites**

**Direct ion-coordination sites**

At the lowest level of ionic densities, several ion contacts are observed around the C=G pairs and the backbone atoms (Figure 4). However, all the contacts do not lead to ordered coordination sites. At a high contour level, densities are visible only around the O6 atom for RNA and around the O6 and N7 atoms for DNA. Thus, while potassium ions contact, at least episodically, each accessible electronegative atom, coordination sites are observed only in the deep/major grooves. On the average, 0.2 K^+ (RNA) and 0.3 K^+ (DNA) form contacts with the O6 atoms while 0.1 K^+ are located close to the N7 atoms (Table 1). This leads to approximately 0.2 (RNA) to 0.3 K^+ (DNA) in contact with the electronegative base atoms. Further, 0.3 K^+ are located close to the phosphate groups. Thus, 0.5 (RNA) to 0.6 K^+ (DNA) are in direct contact with one C=G pair or, on average, about one K^+ per base-paired GpC dinucleotide step, whatever the nucleic acid type. The RMS deviations of these values are between 10% (RNA) and 20% (DNA), which are indicative of large fluctuations (Table 1) resulting from the fact that only certain binding sites are partially occupied during the MD simulation. However, the data indicate unambiguously that the deep/major groove of these duplexes must be considered as an important ion-binding region.

The K^+ in contact with one O6 atom is generally also in contact with the O6 atom of an adjacent base-pair forming an “ion bridge”. Very interestingly, ion bridges are essentially observed for ≈27% (RNA) and ≈40% (DNA) of the GpC steps, while almost no occurrence is observed for CpG

Figure 4. Ion densities calculated around RNA (top) and DNA (bottom) C=G steps. Low-level densities (contact sites) are shown in purple and high-level densities (strong binding sites) are shown in red. On average, there is respectively 0.55 and 0.59 K^+ (Table 1) per RNA and DNA base-pair (purple density), essentially located in the deep/major grooves and around the anionic phosphate oxygen atoms.
steps (Figure 5). This behavior may be due to the different geometrical arrangement of electronegative atoms lining the deep/major grooves. Although, in all four cases shown in Figure 5, the distance between two adjacent G(O6) atoms seems ideal to bind a K\textsuperscript{+}, the protruding of the two electropositive cytosine amino groups for CpG steps prevents the binding of K\textsuperscript{+} into the deep/major grooves. Thus, the MD simulations allow characterization of an important sequence effect related to ion binding.

**Ion-coordination dynamics**

Given the reduced number of ionic contacts to the duplexes, the residence times of the ions were not calculated with respect to a particular coordination site. Rather, the residence times of the ions in the first coordination shell of the central 12 base-pairs of the RNA and DNA duplexes were evaluated. The highest calculated residence times of K\textsuperscript{+} in the first coordination shell of the central 12 base-pairs approach or even exceed the nanosecond range (Figure 6) and for each sequence of 12 base-pairs, at least three ions are present with residence times exceeding 500 ps. The remaining ions display much shorter residence times. It has to be noted that these data are of qualitative character, especially with regard to the ions that display the longest residence times. However, they indicate that residence times in the nanosecond range can occur in the deep/major groove of these duplexes, especially for GpC steps. The bridging position of the ions between two adjacent O6 atoms is surely an important sequence-dependent effect (Figure 5). No chloride ion is observed in the first coordination shells of the duplexes, as was reported for simulations conducted at a 1.0 M NaCl concentration (Feig & Pettitt, 1999b). However, several chloride ions do penetrate the second coordination shell of the double helices by forming ion pairs with positive ions.

**Hydration**

**Water-coordination sites**

The RNA and DNA base-pairs are each surrounded by 22 hydration sites (Figure 7) occupied by an average of 20.8 and 19.2 water molecules (or 21.4 and 19.8 water molecules and K\textsuperscript{+}, respectively (Table 2). Thus, approximately two additional water molecules are found in the RNA primary hydration shell. At a low density level that represents all the possible contact sites (in Figure 7, the densities due to both the K\textsuperscript{+} and the water molecules are presented), it is apparent that the C=G base-pairs are completely surrounded by

| Table 1. Average number of K\textsuperscript{+}/RNA and K\textsuperscript{+}/DNA contacts calculated over the last 1.5 ns of MD for the 12 central base-pairs |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                                 | Deep/major groove | Shallow/minor groove | Backbone and sugar | |
|                                | (G)N7 | (G)O6 | (C)H42 | (G)N3 | (G)H22 | (C)O2 | O2\textsuperscript{-} | OR | OS | O3\textsuperscript{-} | O4\textsuperscript{-} | O5\textsuperscript{-} | Base-pair | All atoms |
| K\textsuperscript{+} (RNA)     | 0.06  | 0.18  | 0.11   | ≈0   | ≈0   | 0.03  | 0.07  | 0.09 | ≈0 | ≈0 | ≈0 | 0.23 | 0.55 |
|                               | (≈0.09) | (≈0.06) | (≈0.04) |        |        | (≈0.02) | (≈0.02) | (≈0.05) |        |        |        |        |
| K\textsuperscript{+} (DNA)     | 0.14  | 0.27  | 0.13   | ≈0   | ≈0   | /     | 0.04  | 0.09 | 0.05 | ≈0 | ≈0 | 0.31 | 0.59 |
|                               | (≈0.20) | (≈0.20) | (≈0.10) |        |        | (≈0.03) |        | (≈0.07) | (≈0.06) |        |        | (≈0.25) | (≈0.13) |

Numbers in parentheses correspond to standard deviations from average values, calculated over four non-overlapping blocks of three consecutive base-pairs.
water molecules, except evidently on the surfaces where stacking contacts occur and on the 5' and 3'-side of the backbone. At a higher density level, the hydration sites appear much less ordered around DNA than around RNA steps, particularly around the more mobile backbone regions. The hydration sites around the base-pairs are, in each case, relatively well defined. For both helices, three sites are found in the deep/major groove in the vicinity of N7, O6, and N4, and three sites are found in the shallow/minor grooves in the vicinity of N3, N2 and O2 atoms. Additionally, for RNA but not for DNA, a hydration site in the proximity of the C5-H group (RNA site 10, Figure 7) and shared with the OR atom of the attached phosphate group can be recognized (Auffinger & Westhof, 1998c). Thus, close to five water molecules occupying six hydration sites are found in the primary hydration shell of a C≡G base-pair (Table 2). The hydration cones, formed by three water molecules around each anionic oxygen atom of the phosphate groups, are easily recognizable in the high-level densities calculated for the RNA molecule. Although present in the DNA oligomers, these hydration cones are less well defined (see the medium-level densities, Figure 7). Hydration cones are also characteristic of the 2'-OH groups of RNA (Figure 7) and account for the presence of an additional hydration site in the shallow groove compared to the DNA minor groove. For DNA, the absence of these hydration sites is balanced by the presence of a hydration site in regard to the O4' atoms (DNA sites 16 and 22, Figure 7).

The hydrogen bonding percentages (HB%) give a finer picture of the contacts established between the solute and the hydration shell (Table 2). They are defined as the total number of hydrogen bonds established during a single trajectory between a particular solute atom and the surrounding water molecules divided by the total number of configurations analyzed (Auffinger & Westhof, 1997a). It appears that in the deep/major grooves (G)N7 atoms establish good hydrogen bond contacts (∼90%), followed by the (C)N4 amino groups (∼80%). The HB% value close to 60 for the DNA (G)O6 atoms, associated with a high standard deviation value, results from the fact that the O6 atoms are also a principal ion-binding site. In the shallow/minor grooves, the (G)N2 amino group is poorly hydrated compared to its neighboring atoms, as reflected by the size of its associated density (see Figure 7). Interestingly, the RNA (G)N3 atom is less well hydrated than the corresponding DNA atom, despite the presence of the 2'-OH group. The 2'-OH appears to be a relatively poor donor and acceptor group, probably due to its rotational dynamics (Auffinger & Westhof, 1997b).

### Table 2. Average number of H2O/RNA and H2O/DNA contacts as well as hydrogen bonding percentages (HB%) calculated over the last 1.5 ns of MD and the 12 central base-pairs (note that, except in the last column of the Table, potassium ion shave not been taken into account)

<table>
<thead>
<tr>
<th>Deep/major groove</th>
<th>Shallow/minor groove</th>
<th>Backbone and sugar</th>
<th>Base-pair</th>
<th>All atoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>(G)N7</td>
<td>(G)O6</td>
<td>(C)H42</td>
<td>(G)N3</td>
<td>(G)H22</td>
</tr>
<tr>
<td>H2O (RNA)</td>
<td>1.1</td>
<td>1.0</td>
<td>1.3</td>
<td>0.8</td>
</tr>
<tr>
<td>H2O (DNA)</td>
<td>(0.1)*</td>
<td>(0.1)</td>
<td>(0.1)</td>
<td>(0.1)</td>
</tr>
<tr>
<td>H2O (RNA)</td>
<td>1.0</td>
<td>0.9</td>
<td>1.2</td>
<td>0.9</td>
</tr>
<tr>
<td>H2O (DNA)</td>
<td>(0.1)</td>
<td>(0.2)</td>
<td>(0.1)</td>
<td>(0.1)</td>
</tr>
</tbody>
</table>

* Numbers in parentheses correspond to standard deviations from average values calculated over four non-overlapping blocks of three consecutive base-pairs.

* Numbers in parentheses correspond to standard deviations from average values.

* These numbers correspond to the exact number of water molecules surrounding a C≡G base-pair (potassium ions have not been counted).

* Numbers in parentheses correspond to the exact number of water molecules and K⁺ surrounding a C≡G base-pair.
The anionic oxygen atoms of the phosphate groups satisfy each of their three hydrogen bond acceptor sites, while the \( \text{O}^3 \), \( \text{O}^4 \), and \( \text{O}^5 \) atoms are poorly hydrated, especially in RNA, where they are involved in \( \text{C}^6/\text{H} \bullet \text{O}^5 \) and \( \text{C}^2-\text{H}^2 \bullet \text{O}^4 \) contacts (Auffinger & Westhof, 1997b, 1998c; Wahl & Sundaralingam, 1997).

**Water-coordination dynamics**

Water molecules establish long-lived hydrogen bond contacts with the RNA and DNA anionic oxygen OR and OS atoms. For both molecules, the characteristic profiles are slightly different. The longest residence times are observed for the (RNA)OR atoms (\( \approx 700 \text{ ps} \)) followed by the (RNA)OS (\( \approx 500 \text{ ps} \)) atoms (Figure 8), while the residence times for the (DNA)OR and (DNA)OS atoms are comparable (\( \approx 400 \text{ ps} \)). Long residence times are observed for the (DNA)O4' atoms, while significantly shorter residence times are observed for the O3' and O5' backbone atoms, and the RNA 2'-OH groups. No marked difference is seen for the RNA and DNA atoms located in the deep/major grooves. Surprisingly, in the shallow/minor grooves, the residence times are much longer in the vicinity of the (DNA)N3 (\( \approx 700 \text{ ps} \), site 20, Figure 7) and (DNA)O2 (up to 1500 ps, site 18, Figure 7) atoms. These long-lived water molecules are not associated with ions, as none of them is observed in the shallow/deep grooves.

**Figure 7.** Water and ion densities showing the 22 hydration sites surrounding the RNA (top) and DNA (middle and bottom) C=G base-pairs (note that the \( \text{K}^+ \) are considered as water molecules, see also Figure 9). Low-level densities (contact sites) are shown in blue (left) and high-level densities are shown in yellow (left and right). For DNA, the high-density patterns (middle) are comparable to those drawn for RNA (top). The medium-density patterns for DNA (bottom) have been drawn in order to distinguish the 22 hydration sites surrounding the C=G base-pairs. Potential hydrogen bonds between hydrophilic groups and potential C-H...O hydrogen bonds are represented by broken lines (right).
Long-lived water bridges

It is useful to characterize the water molecules involved in long-lived water bridges, since those water molecules are susceptible to having the largest structural impact (Westhof, 1993). In the RNA duplex, 17 water molecules are observed to bridge OR anionic oxygen atoms of adjacent phosphate groups with 300 to 400 ps lifetimes (sites 6 and 10, Figure 7). Thus, such long-lived hydration patterns are observed during 15 to 20% of the simulation. A large number of water bridges with significantly shorter lifetimes is also observed. For the B-form DNA duplex, no long-lived water bridge is found between the major groove phosphate groups because the increased phosphate-phosphate interatomic distance does not allow the formation of such hydration patterns (Saenger et al., 1986). However, the exceptionally long residence times of the water molecules associated with the minor groove O2 and, to a lesser extent, N3 atoms (Figure 8) are associated with the formation of (C)O2(n)…W…O4'(n+1) and (G)N3(n)…W…O4'(n+1) bridges. Five (C)O2(n)…W…O4'(n+1) water bridges with lifetimes ran-
ging between 300 and 650 ps and one (G)N3(n)...W...O4′(n + 1) water bridge with a 600 ps lifetime have been observed.

**Calculated ion densities**

It has been noted above that the deep/major groove O6 and N7 atoms are the main K⁺ coordination sites. Obviously, water molecules also occupy these hydration sites and, therefore, the densities shown in Figure 7 result from the cumulative densities due to the presence of water molecules and K⁺ in the vicinity of O6 and N7. In order to visualize the contribution of the potassium ions to the patterns shown in Figure 7, two types of density profiles have been calculated. The first (Figure 9, in yellow) considers all the atoms in contact with the C=G pairs as oxygen atoms of water molecules (Z = 8), the second (Figure 9, in red) takes into account the presence of K⁺ (Z = 18) in the vicinity of O6 and N7. It appears that the calculated densities increase significantly in the vicinity of O6 for the second model.

**Discussion**

**RNA duplexes are more rigid than DNA duplexes**

A large number of MD simulations have been carried out on DNA duplexes using the particle mesh Ewald (PME) summation method (Zichi, 1995; Cheatham & Kollman, 1997; Duan et al., 1997; Young et al., 1997b; Cheatham et al., 1998; Feig & Pettitt, 1998b; Flatters & Lavery, 1998; Lyubartsev & Laaksonen, 1998; Winger et al., 1998; Young & Beveridge, 1998; Sprous et al., 1999; Bevan et al., 2000; MacKerell & Banavali, 2000; Trantirek et al., 2000), while only two studies concerned RNA duplexes (Cheatham & Kollman, 1997; MacKerell & Banavali, 2000). In agreement with 31P NMR experiments (Shindo et al., 1985), Cheatham & Kollman (1997) emphasized the greater rigidity of RNA over DNA duplexes, as reflected by lower RMS deviations from the starting structure, little sugar repuckering and no correlated B₁ to B₁₁ backbone transitions. This tendency is confirmed by our simulations, which show that the r(CG)₁₂ duplex remains very close to standard A-RNA structures.

For DNA duplexes, force-field dependency effects have been reported. For simulations using an older version of the CHARMM force-field (MacKerell et al., 1995), a transition from a B-DNA to an A-DNA form was observed (Yang & Pettitt, 1996), while for simulations performed with the AMBER force-field (Cornell et al., 1995), A to B transitions were described (Cheatham & Kollman, 1996; Cieplak et al., 1997), leading to the conclusion that there is a force-field-dependent DNA polymorphism (Auffinger & Westhof, 1998d). In addition, sequence effects, as discussed by Feig & Pettitt (1998b), may play a significant role, and a certain level of structural heterogeneity is expected. Our simulation of the d(CG)₁₂ sequence reveals a drift toward a form intermediate between A and B, a rather unexpected observation for simulations performed with the AMBER force-field, although experimentally it is known that C=G-containing sequences are rather A-philic compared to A-T-containing sequences (Pilet & Brahms, 1972; Arnott & Selsing, 1974; Saenger, 1984; Minchenkova et al., 1986; Ng et al., 2000; Trantirek et al., 2000). From the theoretical perspective, poly(dG)-poly(dC) sequences were observed to be A-philic, while poly(dA)-poly(dT) sequences were observed to be A-phobic (Cheatham et al., 1998). Thus, the present MD simulations of the r(CG)₁₂ and d(CG)₁₂ structures display a reasonable agreement with our current knowledge of the structure and dynamics of these A and B-form duplexes, and exemplify the experimentally well-documented greater rigidity of RNA compared to DNA duplexes (Hagerman, 1998, 1997). Thus, the AMBER force-field appears, at least partially, to be successful in reproducing several of the sequence-dependent dynamical properties of these duplexes, and the results are in good agreement with MD simulations performed with the CHARMM27 force-field (Foloppe & MacKerell, 2000; MacKerell & Banavali, 2000).
Hydration sites: comparison between experimental and simulation data

The results of the present simulations are in good agreement with studies based on statistical analysis of crystallographic structures at the base-pair level. Schneider & Berman (1995) characterized several well-defined hydration sites in the vicinity of the four A, G, C, and T bases for DNA duplexes in the B-form. These sites (three in the major groove and three in the minor groove associated with a small density in the vicinity of the (C)6 atom) are all well defined in this study (Figure 7). In agreement with the present data, the smallest experimentally characterized and most labile hydration site is located in the middle of the minor groove in front of the (G)2 amino group. The backbone O3' and O5' ester oxygen atoms are not well hydrated, while O4' shares a hydration site with the (G)N3 or (C)O2 minor groove atoms (Schneider et al., 1998). The cones of hydration around each anionic oxygen atom of the phosphate groups (Schneider et al., 1998) are visible only at low-density contour levels indicating that, as a result of the high mobility of the DNA backbone, these sites are not especially well ordered. Only crystallographic studies conducted at a low temperature, which leads to a significant reduction of the atomic mobility, allow detection of precise water peaks around the DNA phosphate groups (Kopka et al., 1983). Given the overall lower mobility of the RNA structure, the hydration sites located around the bases and the backbone atoms are much better defined than those around DNA (Figure 7). The eight hydration sites surrounding the C=G base-pair are clearly identifiable and match very well with those derived from an analysis of high-resolution crystallographic structures (see Figure 1, Auffinger & Westhof, 1998c). The slight difference observed for the O6 site in the deep groove is probably due to the large number of heterogeneous base-pair steps present in the experimental sample reflecting possible sequence-specific dependence. The hydration cones around the 2'-OH groups and observed in earlier studies (Egli et al., 1996; Auffinger & Westhof, 1997b) are also well defined.

In contrast to crystallography, NMR can provide information about the residence times of specific water molecules (Karplus & Faerman, 1994; Kochoyan & Leroy, 1995; Otting & Liepinsh, 1995). Water molecules have essentially been detected in the minor groove of DNA duplexes containing ApT tracts (Kubinec & Wemmer, 1992; Liepinsh et al., 1992, 1994; Jacobson et al., 1996; Denissov et al., 1997; Jóhannesson & Halle, 1998; Phan et al., 1999). Their residence time, involving significant sequence-dependent effects (Liepinsh et al., 1994; Phan et al., 1999), was estimated to exceed the nanosecond at 4°C, reach 0.6 ns at 10°C, and 0.2 ns at 27°C (Denissov et al., 1997; Phan et al., 1999). In the major groove, the observed residence times, in the 200 to 500 ps range at 10°C, are associated with water molecules close to the methyl group of thymine bases not present in the investigated d(CG)12 sequence (Denissov et al., 1997; Lane et al., 1997; Phan et al., 1999), while water molecules contacting the phosphate groups are relatively labile. Given the differences in the investigated sequences, the present MD results are in qualitative agreement with these NMR data. Further, the RNA shallow groove, despite its increased width, was found by NMR as well as in the present simulations, to be more hydrated than the DNA minor groove of an analogous DNA sequence (Conte et al., 1996; Gyi et al., 1998a). This was attributed to the presence of the RNA 2'-OH groups.

Is RNA more hydrated than DNA?

From Figure 7, it is apparent that the hydration patterns around the A-form RNA duplex are better defined than those around the B-form DNA duplex. In all, 22 hydration sites surround the RNA and DNA C=G base-pairs accommodating
20.8 (RNA) and 19.2 (DNA) water molecules as well as 0.5 and 0.6 K$^+$, respectively (Tables 1 and 2). These values are in good agreement with experimental results giving values in the range of 15-20 water molecules per base-pair (Saenger, 1984). However, derived from MD investigations, Feig & Pettitt (1999a) estimated the number of water molecules in the first hydration shell of C=G pairs to be close to 19.3 (A-DNA) and 20.6 (B-DNA), exemplifying the fact that the evaluations of the number of water molecules present in a hydration shell is slightly dependent upon the algorithm used to calculate them (MacKerell & Banavali, 2000).

Yet, despite an identical number of hydration sites, an excess of close to two water molecules is calculated in favor of the RNA structure, allowing us to state that RNA duplexes are more hydrated than DNA duplexes. This value results from a combination of various factors. Among them, the RNA and DNA duplexes adopt different forms (A and B), which modulate their ability to interact with the solvent. The principle of “economy of hydration” (Saenger et al., 1986), which indicates that the more compact A-DNA form needs less hydration than the B-DNA form, cannot be applied in the present case because of the presence of the RNA 2'-OH group. Interestingly, by using the results presented by Feig & Pettitt (1999a), it is possible to propose the rough hydration scale for C=G base-pairs shown in Table 3. Thus, an estimated difference of three water molecules per C=G base-pair between the two RNA and DNA A-form duplexes can be attributed to the presence of the 2'-hydroxyl group. This scale may be useful in the interpretation of experimental data for which precise structural information is not available.

### Short-lived and long-lived ion-binding sites

The main ion-binding sites for the selected RNA and DNA sequences are located on the Hoogsteen edges, close to the (G)O6 and (G)N7 electronegative sites. The occurrence of such ion-binding sites in electronegative pockets in the nucleic acid grooves was proposed earlier on the basis of the observation of the intrusion of a sodium ion into a DNA minor groove (Young et al., 1997a) and is confirmed by the present simulations and those from other groups for RNA (Auffinger & Westhof, 1997a; Cheatham & Kollman, 1997), for DNA (Cheatham & Kollman, 1997; Young et al., 1997b; Lyubartsev & Laaksonen, 1998; Feig & Pettitt, 1999b), and for DNA triple helices (Mohan et al., 1993; Weerasinghe et al., 1995a). Our observations, which indicate the occurrence of an ion bridge between the two deep/major groove (G)O6 atoms of adjacent base-pairs, are consistent with the model proposed by Young et al. (1997a). In addition, the present data demonstrate a significant deep/major groove ion-binding dependence, as almost all observed ion bridges involve GpC and not CpG steps (Figure 5). This finding may be of importance for binding studies trying to understand the affinities of specific amino acids or other chemical groups to particular base-pair steps.

From a dynamical point of view, unlike the ions located in the grooves, which display residence times in the 500 ps range, the ions coming into contact with the phosphate anionic oxygen atoms are not well ordered and do not display residence times exceeding 50 to 100 ps (Figure 6). Interestingly, no ion-binding seems possible in the shallow/minor groove of C=G steps, while binding has been observed by crystallography in the minor groove of ApT steps (Tereshko et al., 1999) exemplifying, along with other MD simulations (Cheatham & Kollman, 1997; Young et al., 1997a; Feig & Pettitt, 1999b), the occurrence of sequence-specific ion-binding patterns. Thus, the present results are consistent with previous propositions (Young et al., 1997a; McFail-Isom et al., 1999) of a hybrid ion/solvent model, indicating that the monovalent ions, present in the surrounding of the nucleic acid, share hydration sites with water molecules. It is noteworthy that the different density maps shown in Figure 9 are indicative of some errors that may be introduced into the interpretation of electronic densities resulting from medium resolution crystallographic studies when one considers that only water molecules are present in the first hydration shell of nucleic acids.

### Differences in gel electrophoretic mobility result from hydration properties

It has been established that RNA duplexes display reduced electrophoretic mobility in polyacrylamide gels compared to DNA duplexes of the same sequence and length, despite possessing the same formal negative charge (Ratmeyer et al., 1994; Lesnik & Freier, 1995; Bonifacio et al., 1997; Hagerman, 1997; Gyi et al., 1998b). This behavior was attributed to either a difference in hydration between the two nucleic acids or to a difference in their ion-binding abilities. Indeed, the ionic atmosphere should be more concentrated around A-form than around B-form structures, as the axial charge density is higher for the former molecules, leading to a higher degree of neutralization of the duplex and a resulting reduced electrophoretic force.

The present results (Tables 1 and 2) indicate that, although there is a significant difference in the hydration of the two duplexes (approximately two

| Table 3. Estimated difference of water molecules in the first hydration shell of A-DNA, B-DNA and A-RNA nucleic acid (CG)$_2$ duplexes |
|----------------|----------------|----------------|
| A-DNA | B-DNA | A-RNA |
| 0 H$_2$O | +1 H$_2$O | +3 H$_2$O |
| 0 H$_2$O | / | +2 H$_2$O |
| 0 H$_2$O | / | / |
water molecules per base-pair in favor of RNA), there is no noticeable variation in the number of counterions in direct contact with the atoms of the nucleic acids, especially in the deep/major groove (Table 1). These ions, which are directly bound to the helices and should therefore migrate with the duplexes, should have the strongest impact on their respective mobility. Thus, it appears that in order to understand the experimentally observed difference in mobility, the two additional water molecules found in the RNA first hydration shell associated with the lower mobility of the solvent molecules surrounding the RNA compared to the DNA molecules should be seriously considered.

About reasonable simulation lengths

Several nucleic acid MD simulations overcoming the present 2 ns range have been published (5 ns Young et al., 1997b; 10 ns Feig & Pettitt, 1998b; Young & Beveridge, 1998; 25 ns Bevan et al., 2000), and this raises the relevant question of the relation between the simulation length and its significance. Although long simulations necessary in order to extensively sample the conformational space of the studied system (Caspard, 1995; Clarage et al., 1995), a few of these nanosecond simulations of nucleic acids sampled infrequent or improbable events (Cieplak et al., 1997; Feig & Pettitt, 1998b). Also, in some instances, it would be desirable to extend the simulation times; indeed the results obtained by several groups indicate that long simulations would sample only protocol and force-field-dependent conformational regions of the "theoretical" phase space, which may not be congruent with the "real" phase space of the studied system (Auffinger & Westhof, 1998b). For example, in some MD simulations, severe and undesirable structural distortion appeared only after running several nanoseconds of simulations (Ravishanker et al., 1997; Cheatham et al., 1999). Such effects have been documented for simulations using several protocols of increased accuracy for the determination of the electrostatic interactions (Auffinger & Westhof, 1998b). It was made clear that sampling the conformational space over long timescales while using inaccurate protocols or inappropriate representations of the environment (Auffinger et al., 1999) is of little use. Beyond doubt, even the best force-fields and MD methods available present some deficiencies that place boundaries to their applicability, despite continual and praiseworthy efforts to improve them (Cornell et al., 1995; Cheatham et al., 1999; Fopolpe & Mackerrell, 2000; Mackerrell & Banavali, 2000). Thus, the reliability of MD simulations has first to be assessed on the shorter time-scales (which extend presently into the nanosecond regime) before exploring the longer time-scales.

Therefore, as an alternative to long MD simulations (Young et al., 1997b; Feig & Pettitt, 1998b; Bevan et al., 2000; Daggett, 2000) and multiple MD simulations (Auffinger et al., 1995; Auffinger & Westhof, 1997a; Caves et al., 1998; Worth et al., 1998), the present strategy of investigating the dynamical behavior of a structure containing highly repetitive motifs (e.g. C=G base-pairs) appears successful in getting interesting and valuable statistical data. This is demonstrated, for example, by the ability of the simulations to discriminate between the ion-binding features of GpC and CpG steps.

Summary and Perspectives

The present MD simulations give, in good agreement with a broad range of experimental data, a precise view of the intrinsic dynamics of an RNA and a DNA duplex, as well as of the hydration shell surrounding the important C=G Watson-Crick base-pairs. The RNA helix is calculated to be less flexible than the DNA duplex, and to remain closer to its starting structure (A-form) than the DNA duplex (B-form). It has been shown that C=G base-pairs are surrounded by 22 hydration sites occupied by \( \approx 21.4 \) (RNA) and \( \approx 19.8 \) (DNA) water molecules and \( K^+ \). This difference of two water molecules can be ascribed to the presence of two hydroxyl groups per RNA base-pair. From the present and previous (Feig & Pettitt, 1999a) results, it is possible to propose the following scale of hydration for C=G base-pairs: \( A\text{-DNA} < RNA \) (\(+3 \text{H}_2\text{O}\)) and \( B\text{-DNA} < RNA \) (\(+2 \text{H}_2\text{O}\)), indicating that RNA is more hydrated than the two \( A \) and \( B \) DNA forms. From a dynamical point of view, the hydration shell is less well defined around the DNA structure than around the RNA structure, due to the higher mobility of the DNA backbone. Water molecules with long residence times are found mainly around the RNA and DNA phosphate groups (0.5 ns range) and some very long-lived water molecules are observed in the DNA minor groove (1 ns range). As expected, some water molecules form long-lived bridges between anionic oxygen atoms of RNA phosphate groups (0.4 ns range) and some form bridges between the O2/N3 and O4' atoms in the DNA minor groove (up to 0.6 ns). Again, water is recognized as an integral part of nucleic acid structures (Westhof, 1988) and, in the present case, the complex interplay between the hydration shell and the presence (RNA) or absence (DNA) of a ribose 2'-hydroxyl group leads to nucleic acid structures with different shapes and dynamical features.

The potassium ion-binding properties are similar for both nucleic acid structures. Approximately, 0.5 \( K^+ \) form direct contact with each C=G pair. The ions form short-lived contacts with the phosphate group and longer-lived contacts with the major groove O6 and N7 atoms. Most of the ions found in the deep/major groove form bridges between two adjacent (G)O6 atoms. However, for both \( B\text{-DNA} \) and RNA duplexes, potassium ions prefer to interact with GpC instead of CpG steps. Ion residence times in the deep/major groove...
approach 0.5 ns. Thus, MD simulations provide a theoretical opportunity to propose reasonable binding sites for the experimentally unobservable or "lost" cations present in the surroundings of the nucleic acids in crystals and solution. The present results may help to rationalize several experimental observations. The variance in hydration calculated in this study should be definitely considered in order to understand the experimentally observed differences in gel electrophoretic mobility of RNA and DNA duplexes of similar sequences, besides the ion-binding properties, which appear to be roughly identical for both nucleic acid structures. Thus, MD simulations provide a good basis for interpreting data originating from NMR or other experimental techniques for which a detailed three-dimensional structure including the environment is often lacking.

The publication of very high-resolution crystallographic structures reaching and even overcoming the 1.0 Å resolution boundary (Masquida et al., 1999; Soler-López et al., 1999) will allow an in-depth comparison between experimental and theoretical results of the hydration patterns around canonical and non-canonical base-pairs. The use of hydration patterns has already been shown to be useful in the refinement of medium or low-resolution crystal structures (Klosterman et al., 1999). Further, as it has been proposed that protein atoms involved in protein/DNA interactions can overlap with DNA hydration sites (Seeman et al., 1976), the characterization of hydration sites should help defining possible protein/nucleic acid or drug/nucleic acid interactions, as proposed earlier for DNA (Woda et al., 1998).

Computational Methods

System setup

Two 2.4 ns MD trajectories of the r(CG)12 and the d(CG)12 helices at constant temperature (298 K) and pressure (1 atm = 101,325 Pa) have been generated with the AMBER 5.0 simulation package (Pearlman et al., 1995; Case et al., 1997). The duplexes were placed in a box containing 72 KCl (Figure 1). The starting A-RNA and B-DNA structures were generated from fiber models by using the NUCGEN module of AMBER. The calculations employed the all-atom force-field described by Cornell et al. (1995). The van der Waals parameters for the ions were extracted from the work of Dang (Cl−, Dang, 1992; K+, Dang, 1995), which calibrated them according to the SPC/E water potential. It has to be noted that the ions displayed a tendency to form clusters with the present water potential, a trend that has been observed in simulations of 1.0 M NaCl aqueous solutions (Degreve & da Silva, 1999). The box dimensions were chosen in order to ensure a ≈12 Å solvation shell around the duplexes. The ions were placed around the nucleic acids on the basis of the electrostatic potential of the solvated system such that no ion was closer than 5 Å to any solute atom or closer than 4 Å to any other ion. Thus, no ion was initially present in the nucleic acid grooves.

MD setup and equilibration procedure

For the treatment of the electrostatic interactions, the particle mesh Ewald (PME) summation method (Darden et al., 1999; Sagui & Darden, 1999) has been used in order to avoid severe artifacts generated by the use of truncation methods (Auffinger & Westhof, 1998b). The charge grid spacing was chosen to be close to 1.0 Å, and a cubic interpolation scheme was used. A 9.0 Å truncation distance was applied to the Lennard-Jones interactions. The trajectory was run with a time-step of 2 fs and shake bond constraints. The non-bonded pair list was updated every ten steps. The equilibration procedure consisted of 100 steps of steepest descent minimization without positional constraints, followed by two 25 ps of MD simulations, the first with fixed solute atoms and ions, the latter with fixed solute atoms and mobile ions. Then, seven 50 ps MD runs were performed with positional constraints on the nucleic acid atoms of 10, 5, 2, 1, 0.5, 0.1 and 0.01 kcal mol−1 Å−2, respectively, yielding a complete equilibration phase of 400 ps (Figure 2). Such an equilibration procedure is akin to procedures where the temperature is slowly raised to the target temperature (Auffinger & Westhof, 1998b). However, with this procedure, water molecules and ions may explore a larger conformational space, as they are always maintained at the target temperature. This is especially important for the ions, which explore their conformational space more slowly than the other components of the system. After the 400 ps equilibration phase, a 2.0 ns production run was performed from which only the last 1.5 ns were used to calculate average properties.

Determination of water and ion-binding sites

The direct ion-coordination and hydration sites were determined by using a procedure first developed by Schneider & Berman for the analysis of the hydration of DNA bases (Schneider & Berman, 1995) and later applied to the hydration of DNA phosphate groups (Schneider et al., 1998) and RNA base-pairs (Auffinger & Westhof, 1998c). Briefly, for the base-pairs, the water molecules or ions that are located at less than 3.5 Å from any heavy atom of the C=G pairs and less than 2.5 Å above or below the base-pair plane were used to construct the water and ion densities (the out-of-plane 2.5 Å criterion has not been used for the backbone atoms). An occupancy factor of 0.5 was ascribed to the ions and water molecules belonging to the coordination shell of two base-pair steps (especially for those close to the (G)O6 atoms in order to
avoid counting them twice during the averaging procedure. The distributions of the water molecules proximate to the base-pairs were transformed into “pseudo-electron” densities through a Fourier transformation by using the SFALL program of the CCP4 library (CCP4, 1994) and the O program (Jones et al., 1991) was used to display the calculated pseudo-electron densities.

The out-of-plane 2.5 Å boundary was set in order to exclude water molecules belonging to the first hydration shell of neighboring base-pairs (the average interbase distance is close to 3.4 Å). The 3.5 Å water-to-base and ion-to-base distance criteria correspond to the first minimum in the corresponding calculated radial distribution functions. Interestingly, this criterion is valid for the ion as well as for the water molecules.

The average structures for the C=G nucleotides shown in Figure 2 were obtained by superposing the heavy base-pair atoms of the 12 central C=G steps for 300 configurations (one configuration every 3 ps) over the last 1.5 ns of the trajectories. For the calculation of the hydrogen bonding percentages, HB%, the following standard hydrogen-bonding criteria were used: $d(H\cdots A) < 2.5$ Å and $\theta(D-H\cdots A) < 135^\circ$, where A stands for a hydrogen acceptor atom and D for a hydrogen donor atom.

Acknowledgments

The authors thank Claude Sauter for his help with the crystallographic software as well as Georges Wipff and Etienne Engler for the use of their MD Draw display program (Engler & Wipff, 1994).

References


composition and sequence. J. Am. Chem. Soc. 120, 6859-6870.


Edited by I. Tinoco

*(Received 28 February 2000; received in revised form 20 May 2000; accepted 22 May 2000)*