Available online at www.sciencedirect.com
SCIENCE DIRECT®
International Journal of Mass Spectrometry 240 (2005) 183–193
International Journal of Mass Spectrometry
www.elsevier.com/locate/ijms

Structural motifs of DNA complexes in the gas phase

Jennifer Gidden, Erin Shammel Baker, Alessandra Ferzoco, Michael T. Bowers*

Department of Chemistry and Biochemistry, University of California, Santa Barbara, CA 93106, USA

Received 11 August 2004; accepted 24 September 2004
Available online 5 November 2004

Abstract

DNA duplexes are known to be quite stable in the condensed phase but recent mass spectrometry results have shown that DNA complexes are also stable (at least for a limited time) in the gas phase. However, very little is known about the overall shape of the complexes in a solvent-free environment and what factors influence that shape. In this article, we present recent ion mobility and molecular modeling results that address some issues concerning the gas-phase conformations of DNA duplexes. Examples include the effect of metal ions on Watson–Crick base pairing, investigating the onset of helicity in duplexes as a function of strand length, comparison of the stability of C·G and A·T base pairs, and examining the formation of quadruplex structures.

Keywords: Ion mobility; DNA; Helix; Conformations

1. Introduction

DNA duplexes are stabilized by a number of factors but hydrogen bonding between bases on the two strands and base stacking within each strand are the major contributors [1]. Solvent is believed to be just as crucial to the stability of the duplex because it can provide screening of the negatively charged phosphate backbones [2]. Thus, most structural and characterization studies of DNA are performed in the condensed phase and obtaining gas-phase data has been believed to be nearly impossible as an increase in charge repulsion from the absence of solvent screening and the reduced favorability of base stacking should significantly destabilize the duplex structure [2].

However, in 1993 Ganem et al. [3] and Light-Wahl et al. [4] demonstrated that DNA duplexes could be successfully transferred, intact, from solution to the gas phase using electrospray ionization mass spectrometry (ESI-MS) [5]. A year later, Doktycz et al. showed that DNA duplexes could survive in the gas phase for hundreds of milliseconds in a quadrupole ion trap [6]. Since that time, a number of papers and review articles have been written describing the usefulness of mass spectrometry in characterizing DNA duplexes in the gas phase [7–10]. The types of results gathered from these studies range from the determination of base composition and sequence of DNA strands to the identification of ligand binding sites and post-translational modification sites. Despite this wealth of data, the overall gas-phase structures of the duplexes and the factors that influence these structures and the interactions between the two strands are not well understood (not to mention whether the gas-phase structures are representative of their solution phase counterparts or undergo major conformational changes).

Several groups have attempted to address these issues using selective dissociation of the duplexes and have shown that specific hydrogen bonding and base stacking may, indeed, be conserved in the transfer from solution to the gas phase. Gabelica and DePauw, for example, used collision-induced dissociation (CID) to examine a series of 16-mer duplexes [11,12]. They observed that the relative kinetic stabilities of the gas-phase duplexes correlated well with relative stabilities in solution measured by thermal denaturation (monitored by UV spectrometry). CID fragmentation yields also paralleled calculated solution melting enthalpies. They attributed these results to the retention of hydrogen bonding

* Corresponding author. Tel.: +1 805 893 2893; fax: +1 805 893 8703.
E-mail address: bowers@chem.ucsb.edu (M.T. Bowers).

1387-3806/$ – see front matter
doi:10.1016/j.ijms.2004.09.011

and base stacking interactions in the gas phase that are present in solution. Schnier et al. studied the kinetics of dissociation of several complimentary and non-complimentary 4- to 7-mer duplexes using blackbody infrared radiative dissociation (BIRD) [13]. They observed that activation energies for dissociation of complimentary duplexes were higher than those of non-complimentary duplexes and these activation energies correlated with solution dimerization enthalpies, thus, indicating that Watson–Crick pairing was conserved in the gas phase (corroborated by molecular dynamics calculations). Griffey et al. examined the CID fragmentation of 8- to 14-mer duplexes with mismatched base pairs and observed preferential cleavage at the site of the mismatch (suggesting that Watson–Crick pairing was preserved in the rest of the duplex) [14].

Although these studies indicate that DNA duplexes conserve a portion of their solution phase character in the gas phase, their overall conformation in the gas phase remains a major question. In solution, DNA duplexes are often helical, taking on A-, B-, or Z-forms, but very little is known about whether these helices exist in the gas phase and what factors influence them. In this article we will report on some of our recent ion mobility and molecular modeling studies that have focused on these issues and examine factors such as how metal ions affect Watson–Crick pairing and the importance of strand length and sequence on the overall conformation of the duplex.

2. Experimental and theoretical methods

2.1. Ion mobility measurements

The mobility of an ion (K) is simply a measure of how fast the ion drifts through a buffer gas (ν_d) under the influence of a weak, uniform electric field (E) [15].

$$\nu_d = K \cdot E \tag{1}$$

For large ions, the mobility also depends significantly on the geometric shape of the ion. Compact ions with small collision cross-sections undergo fewer collisions with the buffer gas and hence drift faster than more extended ions. Thus, Eq. (1) can be re-written as

$$t_d = \frac{C_1}{K_o} = C_2\sigma \tag{2}$$

where t_d is the drift time of the ion; K_o, the reduced mobility (scaled to standard state); σ, is the ion's collision cross-section, and C_1 and C_2 are given below [15].

$$C_1 = l^2 \cdot \frac{273p}{760T} \cdot \frac{1}{V} \tag{3a}$$

$$C_2 = \frac{16N_o}{3e}\left(\frac{2\pi}{\mu k_b T}\right)^{-1/2} \tag{3b}$$

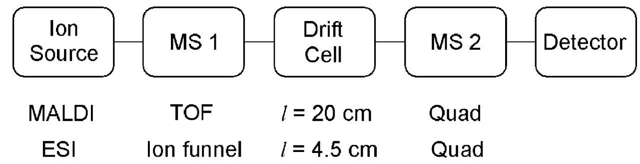

Fig. 1. Schematic of the basic ion mobility/mass spectrometry instrumental setup.

In the above equations, l is the drift length; p, the pressure of the buffer gas; T, temperature; V, the drift voltage applied across the cell; N_o, the number density of the buffer gas at STP; e, the charge of the ion; μ, the reduced mass of the ion and buffer gas; and k_b, Boltzmann's constant.

A schematic of the experimental setup used to measure ion mobilities is shown in Fig. 1 [16–19]. Ions are generated by electrospray ionization (ESI) or matrix-assisted laser desorption/ionization (MALDI) [20], mass selected with the first mass analyzer, and injected into a drift cell containing helium. For the ESI experiments, the continuous ion beam from the source is gated so that the ions can be pulsed into the drift cell, thus, triggering a timer. For the MALDI experiments (which has a pulsed ion beam), the laser pulse is used to trigger the timer. The ions injected into the drift cell are quickly thermalized by collisions with helium ($\sim10^5$ collisions) and drift through the gas at constant velocity under the influence of a weak electric field. Ions exiting the cell are mass analyzed with a quadrupole mass filter and detected as a function of time on a multi-channel scalar, yielding an arrival time distribution (ATD).

Arrival times at the detector, t_a, are the sum of t_d and t_o, where t_o is the time the ion cloud spends outside the drift cell before reaching the detector. Hence,

$$t_a = t_d + t_o = l^2 \cdot \frac{273p}{760T} \cdot \frac{1}{V} \cdot \frac{1}{K_o} + t_o \tag{4}$$

and a plot of t_a versus p/V should yield a straight line with a slope inversely proportional to the mobility of the ion and an intercept equal to t_o.

2.2. Theoretical modeling

Conformational information about the ions is obtained by comparing the cross-sections obtained from the ATDs to calculated values of theoretical models. The AMBER 7 [21] suite of molecular dynamics software (using the Amber99 force field) as well as DFT calculations (using the B3LYP hybrid functional and the LACVP basis set) [22,23] calculations were used to generate candidate structures of the duplexes and their corresponding collision cross-sections were calculated using an angle-averaged projection model [24,25] and a hard sphere scattering model [26]. In the molecular dynamics calculations, hundreds of low energy structures are generated via a simulated annealing method that has successfully predicted low energy structures for numerous biological and synthetic polymers [18,27–32]. The average cross-section of the lowest 5–10 kcal/mol structures (which usually have only minor structural differences) is then compared to experiment for conformational identification. In the DFT calculations, only one low energy structure is obtained for a given starting geometry and its cross-section is calculated 10 times and averaged for comparison to experiment. Starting geometries of the DNA duplexes were obtained using the NUCGEN utility in AMBER 7 as well as published X-ray structures.

2.3. Sample preparation

All DNA samples were purchased from Sigma-Genosys (The Woodlands, TX) and used without further purification. The samples were annealed in H_2O or NH_4OAc with duplex concentrations ranging from 30 to 300 μM and NH_4OAc concentrations ranging from 20 to 150 mM. In the MALDI experiments (metal-dinucleotide complexes), 2,5-dihydroxybenzoic acid was used as the matrix and a water/methanol mixture as the solvent (see [32] for full procedure). In the ESI experiments, the annealed DNA solutions were diluted to 30–75 μM in H_2O and sprayed in a 98:2 mixture of NH_4OAC/NH_4OH or a 49:49:2 mixture of $H_2O/MeOH/NH_4OH$.

3. Applications

3.1. Metal ions and Watson–Crick bonding

The importance of metal cations interacting with DNA was first realized in the 1920s when Hammarsten reported on the need for metal cations to be present in cells to help neutralize the overall negative charge on DNA [33]. However, metal-DNA studies did not really begin in earnest until the late 1960s after Rosenberg and co-workers discovered that cisplatin (*cis*-$[Pt(NH_3)_2Cl_2]$) was an effective antitumor agent and ensuing work suggested that the binding of Pt to DNA bases was largely responsible [34–37]. In recent years, a major focus of metal-DNA studies has been identifying the role metal cations play in stabilizing quadruplex structures [38–41].

Metal cations are usually found near the negatively charged phosphate groups on the DNA backbone but numerous studies have shown that metals can bind almost anywhere on the DNA molecule [42]. The next most popular sites are the nucleobases and recent ab initio calculations have even indicated that the proper placement of metal cations on nucleobases can actually enhance Watson–Crick bonding between complimentary pairs [43,44]. One particular site that has generated recent interest is the O6 atom on guanine bases. It is at this site that metal cations such as Na^+ or K^+ bind, stabilizing quadruplex structures [38–41].

MALDI-TOF spectra obtained in our lab on a series of dinucleotides cationized by Na^+ indicated that up to 7 Na^+ could bind to a given duplex and up to 10 Na^+ could bind to triplexes, with n Na^+ ions replacing $(n-1)$ hydrogens [32]. Ion mobility and molecular modeling results indicated

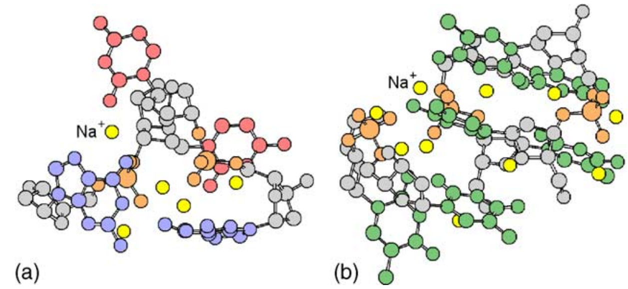

Fig. 2. Lowest energy structures of (a) $[(dCG\cdot dCG)+5Na-4H]^+$ and (b) $[(dTT\cdot dTT\cdot dTT)+10Na-9H]^+$ obtained by molecular mechanics calculations. Cytosine is shown in red, guanine in blue, and thymine in green. Na^+ ions are yellow and the phosphate groups are orange. The Na^+ ions bind preferentially to oxygen atoms on the phosphate groups and bases.

that the Na^+ ions initially cluster around the deprotonated phosphates, but as more Na^+ are added, they become more dispersed and bind to multiple sites and groups in the duplex/triplex. In order to accommodate all of the Na^+ ions and maintain the +1 charge state commonly observed in MALDI experiments, the bases must deprotonate along with the phosphate groups. Thymine and guanine are the only good candidates and can deprotonate at N3 and N1, respectively. However, the Na^+ ions do not necessarily bind to those particular deprotonated sites, preferring instead to coordinate to oxygen atoms. Examples of duplex and triplex structures with multiple Na^+ ions are shown in Fig. 2.

Fig. 3 shows arrival time distributions (ATDs) for a dCG·dCG duplex cationized by a variety of alkali, alkaline, and transition metal ions [45]. Single, symmetric peaks are observed in the ATDs for all the metal-duplex ions (Fig. 3a), except those cationized by Cu^+, Ag^+, $[Zn^{2+}–H]^+$, and $[Cd^{2+}–H]^+$. In those cases (Fig. 3b), two peaks are present, indicating two distinct isomers exist that have significantly different collision cross-sections (and hence geometries). The cross-sections extracted from the shortest-time peaks in Fig. 3b (229 ± 2 Å^2) are comparable to those determined from the ATD peaks in Fig. 3a but the cross-sections extracted from the longest-time peaks in Fig. 3b are ∼12% larger (255 ± 3 Å^2).

The lowest energy structures determined for the Na^+[dCG·dCG] and Cu^+[dCG·dCG] duplexes are shown in Fig. 4. In the sodiated duplex (Fig. 4a), the Na^+ ion binds to carbonyl oxygens on all four bases, disrupting the Watson–Crick hydrogen bonding between cytosine and guanine. In the modeling (using the AMBER 7 program), the Na^+-duplex was initially placed in a “Watson–Crick” geometry but it quickly rearranged into the more compact structure shown in Fig. 4a. The calculated cross-section of this structure (228 ± 2 Å^2) agrees very well with experimental values. Molecular dynamics calculations on the other alkali and alkaline metal-duplex ions yield similar structures.

DFT calculations (B3LYP/LACVP) [22,23] were used to generate optimized structures for the Cu^+-duplex ions using Watson–Crick geometries and the “globular” geometry

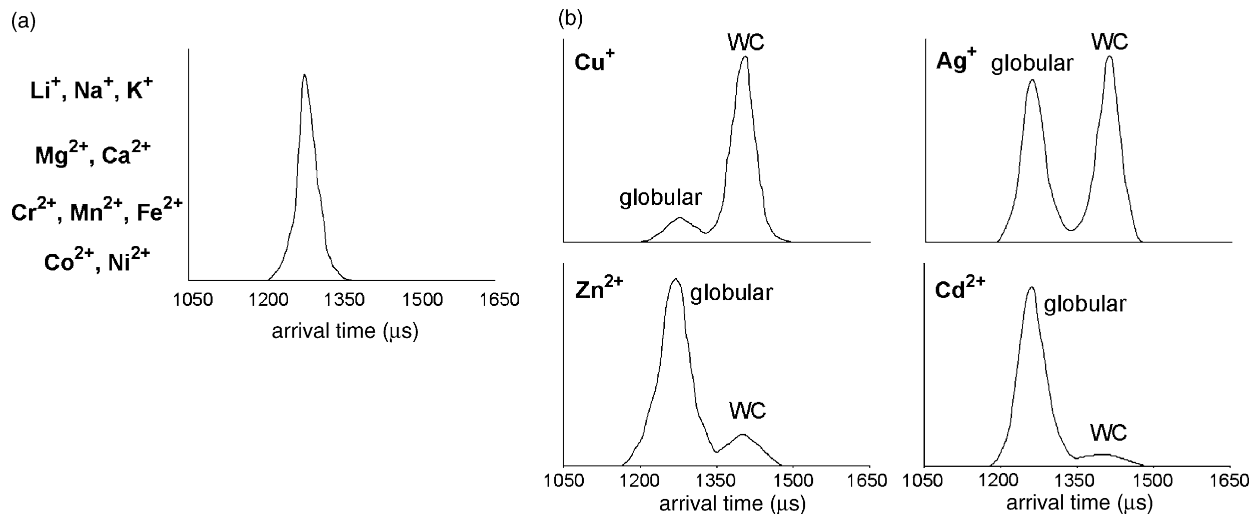

Fig. 3. Arrival time distributions (ATDs) of (dCG·dCG) cationized by different alkali, alkaline, and transition metals (generated by MALDI). Two peaks, corresponding to globular and Watson–Crick structures, are observed only for the d^{10} metals.

shown in Fig. 4a as starting structures. Cu^+ was added to both duplex conformers at many different potential binding sites. In each case, the starting geometries remained intact throughout the optimization procedure and the energies of the final structures of each form were within 1 kcal/mol of each other. In the globular structure, the Cu^+ ion is in a similar position as the Na^+ ion, coordinating to the carbonyl oxygen on each of the four bases. The cross-section of this structure agrees well (within 2%) with the value extracted from the fastest-time peak in the ATD. In the "Watson–Crick" geometry, the Cu^+ ion binds to carbonyl oxygens on the cytosine bases and to the NH_2 groups on the guanine bases. Remarkably, the location of the Cu^+ ion between the bases does not disrupt the W–C hydrogen bonding between cytosine and guanine and the resulting structure has an average cross-section that agrees well with the value obtained from the longest-time peak in the ATD. If the Cu^+ ion is positioned near N7 and O6, the preferred site on free base pairs [42,43], the resulting structure (although a Watson–Crick geometry) is ∼10 kcal/mol higher in energy.

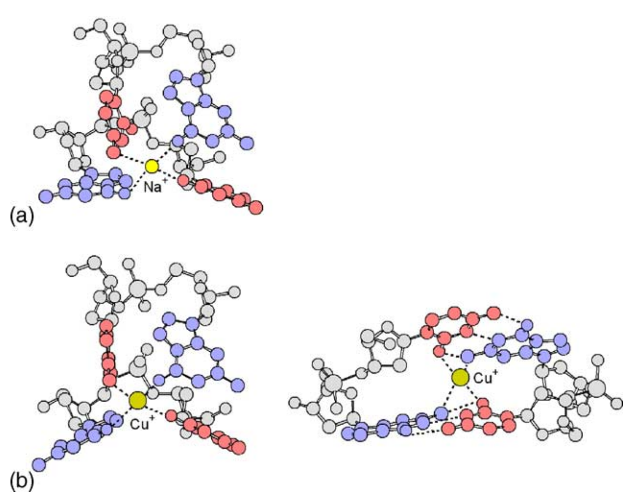

Fig. 4. (a) Lowest energy structure for Na^+(dCG·dCG) obtained by molecular mechanics calculations. Na^+ binds to carbonyl oxygens on all four nucleobases. (b) Lowest energy structures for Cu^+(dCG·dCG) obtained by DFT calculations. In the Watson–Crick structure, Cu^+ binds to O2 on the cytosine bases and N2 on the guanine bases.

Watson–Crick structures are observed experimentally only for dCG·dCG duplexes cationized by d^{10} metals. This result is not well understood as similar ion mobility studies on dAT·dAT duplexes yielded no Watson–Crick structures, regardless of the metal cation. One possible explanation may lie in the fact that these four d^{10} metal cations can also be classified as soft acids (although Zn^{2+} is a borderline acid) whereas all of the other metal cations are hard acids or borderline acids [46,47]. Hard acids tend to form complexes that are dominated by electrostatic interactions but soft acids form complexes with more covalent character. Soft acids also tend to prefer to bind to nitrogen atoms rather than oxygen atoms. In the Watson–Crick structure shown in Fig. 4b, the Cu^+ ion binds to the N2 atoms on guanine and appears to help bridge the NH_2 group on guanine and the O2 atoms on cytosine, keeping the Watson–Crick hydrogen bonding intact. If the metal ion has a high preference for binding to oxygen atoms (like Na^+), the bases must rearrange to accommodate the metal ion and thus, disrupt the hydrogen bonding between the bases.

3.2. Size effects—onset of helicity

In the condensed phase, DNA duplexes can adopt a number of different structures but are commonly found in a double helix arrangement. X-ray studies of DNA fibers have revealed three basic conformations of the double helix (shown in Fig. 5): the B-form, which is stable at high humidity (∼92%) [48], the A-form, which is dominant at lower hu-

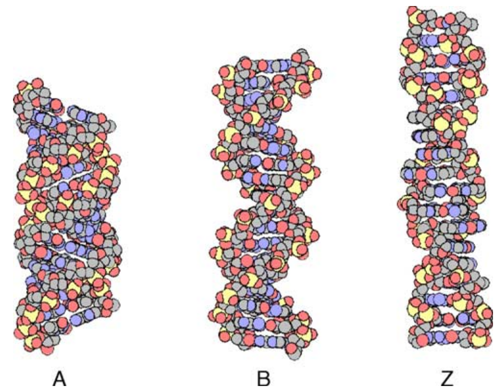

Fig. 5. Canonical structures of the three main types of helical duplexes: A-, B-, and Z-DNA.

midity (∼75%) [49], and the Z-form, which has been found for $(dCG)_n$ repeats at high salt concentrations [50]. As described previously, mass spectrometry studies have shown that DNA duplexes can survive in the gas phase and conserve at least some of the Watson–Crick hydrogen bonding and base stacking interactions that are present in the X-ray data. However, very little is known about the overall conformation of the duplexes in the gas phase and whether they retain any helical character.

In this section, we briefly describe some recent ion mobility measurements designed to answer some questions about the gas-phase conformations of DNA duplexes and in particular address how the length of the duplex affects its shape [51]. Fig. 6 shows arrival time distributions for a series of deprotonated $d(CG)_n{\cdot}d(CG)_n$ duplexes generated by electrospray ionization. ATDs are shown for the 6-mer ($n = 3$) to the 18-mer ($n = 9$) and indicate a dramatic change in structure at the 8-mer length. The ATD for the 8-mer is the only one with two peaks, indicating two distinct isomers exist, and the time difference between the two peaks corresponds to a difference in cross-section of 130 Å^2 (a 25% change!).

Theoretical structures assigned to each $d(CG)_n{\cdot}d(CG)_n$ duplex, based on molecular modeling (AMBER 7) and comparison of experimental and theoretical cross-sections, are also shown in Fig. 6. The 4-mer (not shown) and 6-mer duplexes are globular with only one C·G pair hydrogen bonded in a Watson–Crick arrangement. The fastest-time peak in the 8-mer ATD can also be assigned to a globular structure (with three Watson–Crick pairs) but the longest-time peak in the ATD can only be assigned to a more extended structure.

In the theoretical modeling of the 8-mer, the duplex was initially arranged in a canonical A, B, or Z helix. However, during the 700 K simulated annealing procedure used to generate low-energy candidate structures [30], all three helices collapsed into lower energy globular forms. If the 8-mer duplex is placed in a water box, on the other hand, it will remain helical. When the water box is removed, the helical structure will become distorted but will not immediately collapse into a globular conformation. In fact, the 8-mer duplex will remain in a quasi-helical structure (retaining seven of eight W–C

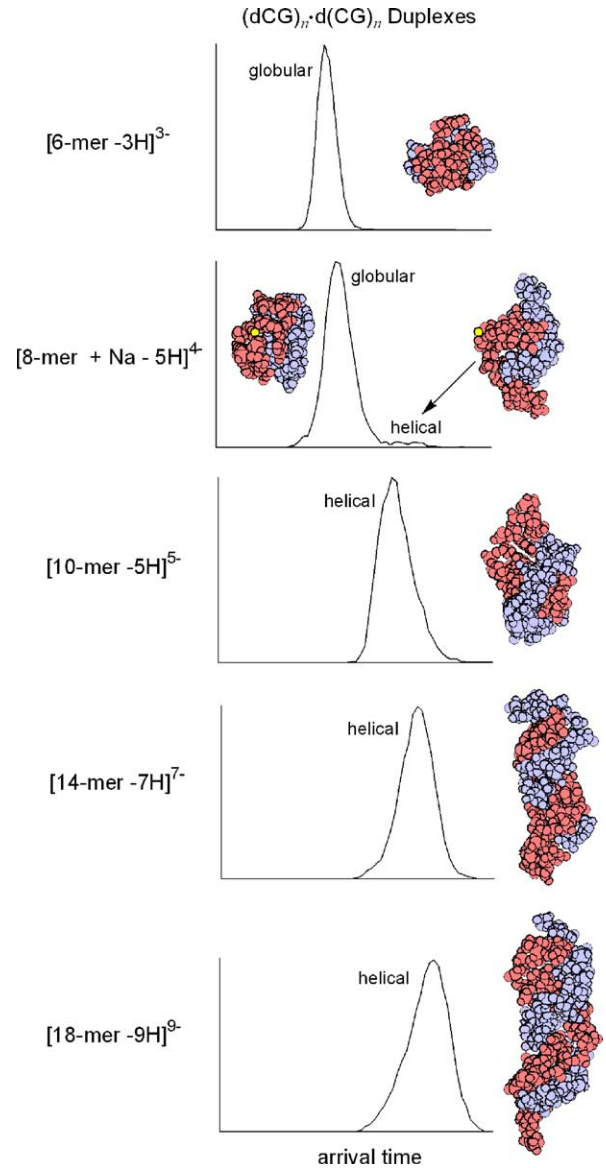

Fig. 6. ATDs and theoretical structures of $d(CG)_n{\cdot}d(CG)_n$ for $n = 3 - 9$. The smallest duplexes are globular but the larger duplexes have helical structures. The transition point appears to be at $n = 4$ (8-mer length) where both globular and helical structures are observed in the ATD.

pairs) throughout 2 ns of 300 K dynamics and the average cross-sections of the final structures match the experimental value extracted from the longest-time peak in the ATD. Similar results are obtained for the 10-mer through 18-mer duplexes and examples of the final helical structures from the 300 K dynamics are shown in Fig. 6 (starting with the A helix).

One ongoing question in ESI-MS studies is whether the structures of biological molecules in the gas phase are representative of their condensed phase counterparts, allowing a relationship to be drawn between solvent-free measurements and those performed in solution. In the $d(CG)_n{\cdot}d(CG)_n$ series, molecular modeling predicts that globular conformers are the lowest energy gas-phase structures for all of the duplexes but the ATDs clearly indicate helical structures are present for the larger systems. This can only be accounted for if the larger duplexes were originally helical in solution and retained those "solution phase" structures for a limited time (at least on the experimental time scale of 1 ms) in the gas phase. The gas-phase conformations of the duplexes are not textbook A-, B-, or Z-DNA helices (and they may not have been that way in solution), but they certainly retain helical structures and preserve a high percentage of Watson–Crick pairs in the gas phase. Thus, the structures observed in the ion mobility data are most likely metastable helical structures. As the length of the duplex increases, the time it takes for the helical structure to relax and rearrange into its lowest-energy globular form most likely increases as well. For example, the helical structure is a minor component in the ATD for the 8-mer duplex but becomes the dominant form in the ATDs of the 10-mer through 18-mer duplexes. This process is shown schematically in Fig. 7.

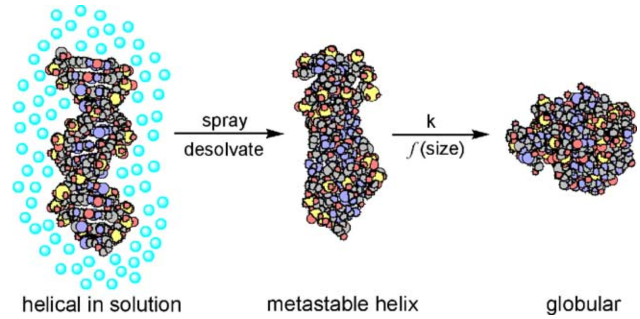

Fig. 7. Possible schematic of events that explains the results obtained for the longer $d(CG)_n{\cdot}d(CG)_n$ duplexes ($n = 4 - 9$). The duplexes are helical in solution and retain that shape (with minor changes) when desolvated. These helical structures will eventually collapse into the lowest energy globular forms, but as the length of the duplex increases, the time it takes for the collapse to occur also increases.

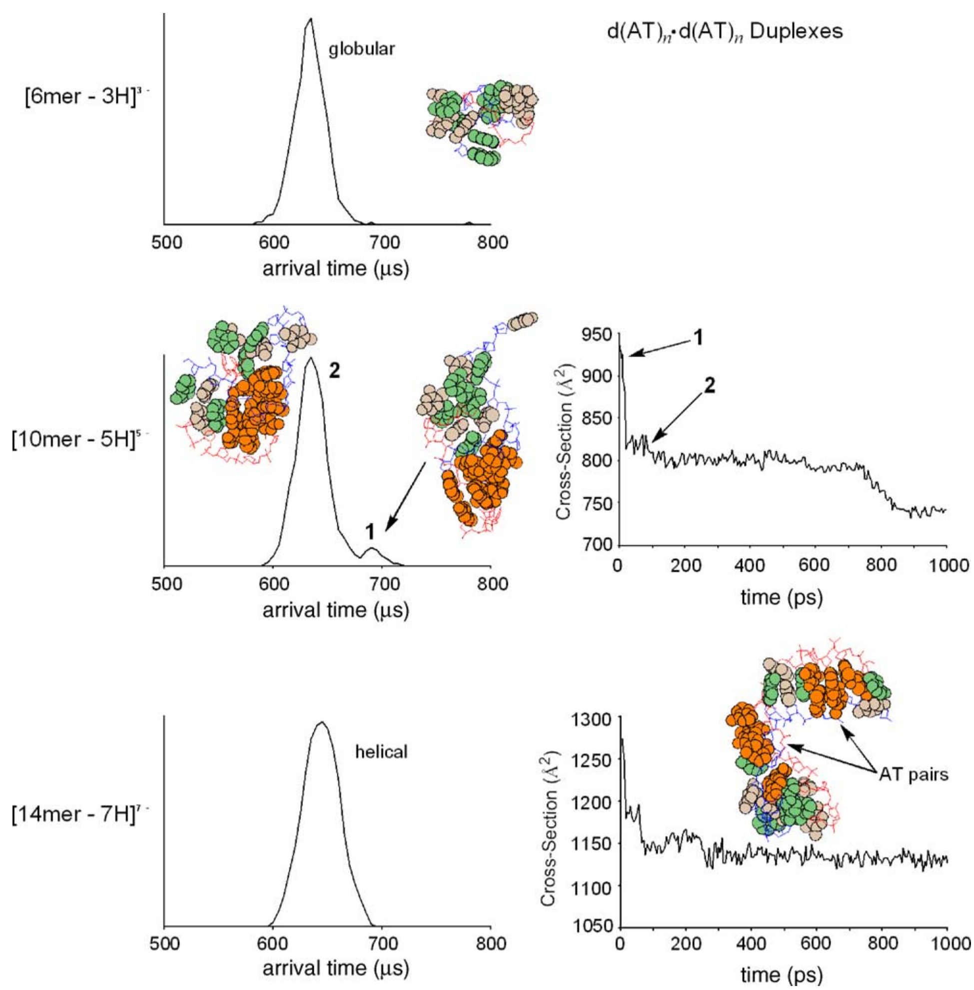

Fig. 8. ATDs and theoretical structures for $d(AT)_n{\cdot}d(AT)_n$ for $n = 3 - 7$. Watson–Crick A·T pairs are shown in orange. Unpaired adenines are green and unpaired thymines are brown. Also shown are cross-section vs. dynamics time plots for the 10- and 14-mer. Structures were saved every 5 ps and their cross-sections calculated.

Fig. 8 shows ATDs and theoretical structures for a series of deprotonated $d(AT)_n \cdot d(AT)_n$ duplexes generated by electrospray ionization. Like the $d(CG)_n$ series, the smallest $d(AT)_n$ duplexes are globular whereas the longer duplexes retain helical structures. However, several significant differences exist between the $d(CG)_n$ and $d(AT)_n$ series. First, the percentage of A·T pairs that remains in a Watson–Crick arrangement (shown in orange in Fig. 8) is significantly smaller than that observed for the C·G pairs. In the latter case, 85–90% of the C·G pairs remained in a Watson–Crick arrangement after 2 ns of 300 K dynamics (for the 8-mer through 18-mer). That number drops to 50% for the d(AT) 10- and 14-mer duplexes.

The second difference centers around the transition point from globular to helical structures, shifting from $n = 4$ in the $d(CG)_n$ series to $n = 5$ in the $d(AT)_n$ series. The ATD for the d(AT) 8-mer duplex could not be attained [51] but the ATD for the 10-mer shows two distinct isomers are present (like the d(CG) 8-mer) and molecular modeling indicates the d(AT) 10-mer is not as stable in the gas phase as in the d(CG) case. A plot of cross-sections versus dynamics time for the d(AT) 10-mer is shown in Fig. 8. Structures were saved every 1 ps for the first 10 ps and every 5 ps after that. Starting with a canonical B-DNA helix, the 10-mer immediately (before 1 ps) converts into structure **1**, which is tentatively assigned to the longest-time peak in the ATD based on comparison of experimental and theoretical cross-sections. Structure **1** is not stable and quickly converts into structure **2**, which is assigned to the fastest-time peak in the ATD, after ∼20 ps of dynamics. The main difference between the two structures is that the unpaired A and T bases begin to fold around the duplex and stack differently, forming a less helical structure. After ∼750 ps, theory predicts structure **2** will collapse into a more compact structure (breaking two more A·T pairs) but this final structure is not observed in the ATD. In any case, the d(AT) 10-mer is clearly not as stable in a helical arrangement as the d(CG) 10-mer (which retains 9 of 10 Watson–Crick pairs). The d(AT) 14-mer duplex will remain helical throughout 1000 ps of 300 K dynamics (the final structure at the end of the dynamics run is shown in Fig. 8) but it only retains 7 of 14 A·T pairs and has a significant bend in the middle that is not observed for the d(CG) 14-mer duplex.

3.3. Sequence effects (CG versus AT)

In reality, DNA duplexes are not likely to be in an exact A-DNA or B-DNA form, even in the condensed phase. One reason is that the base composition and sequence of the duplex can induce local variations in structure [1,2]. Long runs of adenine, for example, can cause the DNA duplex to bend and extended runs of guanine bases in DNA can lead to triplex or quadruplex formation. C·G runs have been shown to induce local A-form structures in an otherwise B-form DNA duplex and A + T rich sequences are known to actually form stable unwound structures. The overall base sequence can also have a dramatic result on the structure of the DNA duplex. Inverted repeat sequences can form cruciform structures, mirror repeat sequences are common in triplex structures, and direct repeat sequences can lead to mispaired DNA structures.

In the previous section, it was shown that DNA duplexes were not only stable in the gas phase; they could retain helical structures provided the length of the duplex chain was sufficient. The results also showed, however, that hydrogen bonding and base stacking were better preserved in the C·G duplexes than the A·T duplexes. This is not a surprising result as DNA melting temperatures and denaturation curves have long indicated that C·G pairs are more difficult to disrupt than A·T pairs [2,12]. In this section, we further explore the differences between C·G and A·T base pairs and focus on how the sequence of the duplex affects its overall gas-phase conformation.

Fig. 9 shows ATDs and theoretical structures of a series of 14-mer duplexes (with 7^- charge states) containing A·T

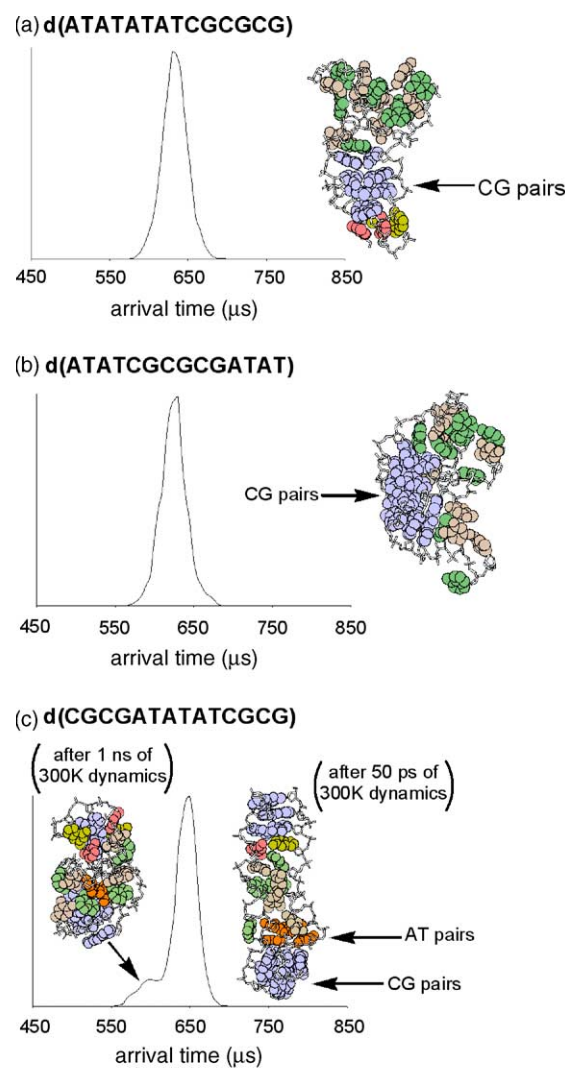

Fig. 9. ATDs and theoretical structures for a series of 14-mer duplexes with AT and CG groups in different sequences. In each case, more C·G pairs (blue) are retained than A·T pairs (orange). No A·T pairs remain intact if they are placed on the ends of the duplex. Unpaired adenine is green, thymine is brown, cytosine is red, and guanine is gold.

and C·G pairs in different sequences. In each case, more C·G pairs are retained (shown in blue) than A·T pairs (shown in orange). The A·T pairs appear to more easily disrupted when they are placed at the ends of the duplex as neither the d(ATATATATCGCGCG) nor d(ATATCGCGCGATAT) duplex have any A·T Watson–Crick pairs left at the end of 1 ns of 300 K dynamics. If the A·T pairs are placed in the middle of the duplex [i.e., d(CGCGATATATCGCG), Fig. 9c] two A·T pairs are retained and the overall structure becomes more elongated, resembling d(CGCGAATTCGCG) structures generated by molecular dynamics calculations [52]. This elongated structure, assigned to the longest-time peak in the ATD, is quite stable (lasting at least 1 ns in the 300 K dynamics) but eventually collapses into a more compact form (in which an additional C·G pair is broken) that is assigned to the fastest-time peak in the ATD.

Shown in Fig. 10 are the measured ATD and theoretical structures for the 5^- charge state of a $d(GA)_5 \cdot d(TC)_5$ 10-mer duplex generated by electrospray ionization. This particular sequence has been shown to form a variety of different intramolecular triplex structures [53–55] but also contains A·T and C·G base pairs, so that comparisons can be made between the two types of Watson–Crick bonding. The lowest energy gas-phase structures of $d(GA)_5 \cdot d(CT)_5$ are globular but the average cross-sections of these structures are significantly smaller than the experimental value. During 300 K dynamics simulations, the duplex remains in a helical arrangement for at least 1 ns and the final structures (starting with A- and B-DNA helices) are shown in Fig. 10. The A-form agrees best with experiment, but in both cases all five C·G pairs (shown in blue in the figure) remain intact whereas only two of the A·T pairs (shown in orange) are still bound to each other at the end of the dynamics run. One interesting structural feature is that the unpaired A and T bases will interact with the C·G pairs. This is most notable at the end of the duplex where an Ade base hydrogen bonds to Gua (NH_2–N7) and a Thy base hydrogen bonds to Cyt (O2–NH_2) on the end C·G pair. Both A and T bases are weakly bound to the C·G pair, forming only one hydrogen bond between it and the base pair, but nonetheless are the beginnings of the formation of an intramolecular triplex-like structure.

3.4. Quadruplex formation

In 1988 Sen and Gilbert reported that a DNA sequence rich in guanine could form stable four-stranded structures, presumably held together by Hoogsteen bonding between the guanine bases [56]. Since that time, a number of NMR and X-ray studies have verified this quadruplex structure, showing a planar array of G-quartets stacked on top of each other [2,36–39]. More recently, ESI mass spectrometry studies have shown that quadruplexes can exist in the gas phase. Several papers have reported "magic number" quartet adducts of guanine and guanosine [57–59], Goodlett et al. have observed $d(CGCG_5CG)$ tetramers [60], and Rosu et al. have detected tetramers of d(TGGGGT) in ESI mass spectra [61].

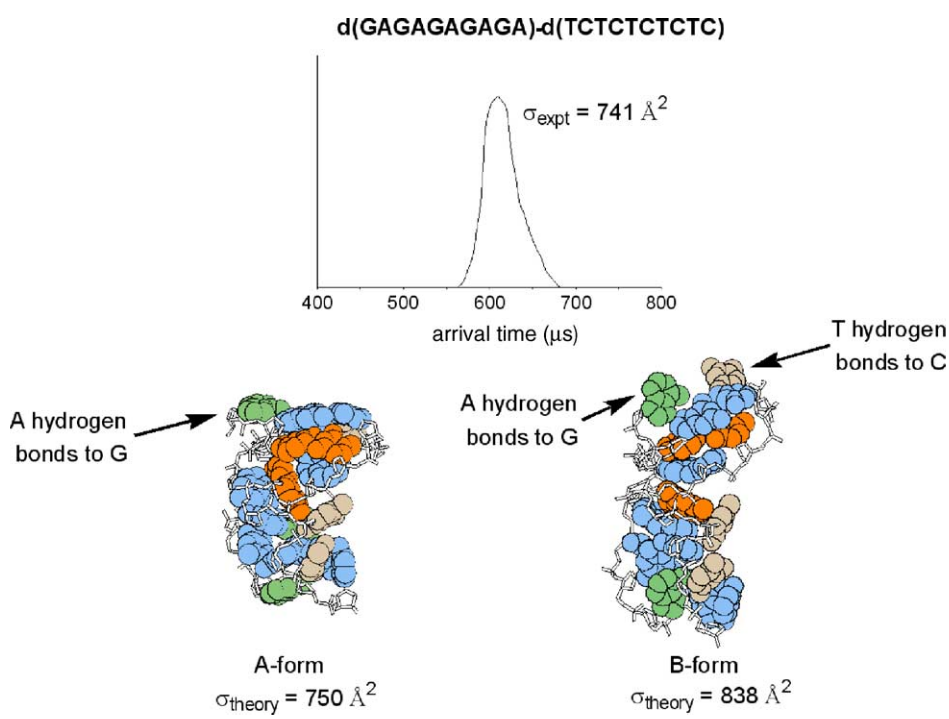

Fig. 10. ATDs and theoretical structures for $[d(GA)_5 \cdot d(CT)_5 - 5H]^{5-}$. The structures shown are the final structures obtained from 2 ns of 300 K dynamics starting with canonical A- and B-DNA geometries. C·G pairs are shown in blue and A·T pairs are shown in orange. Unpaired adenines are green and unpaired thymines are brown. The unpaired A and T bases interact with the C·G pairs, even forming triplex- and quadruplex-like structures at the end of the duplex (see text).

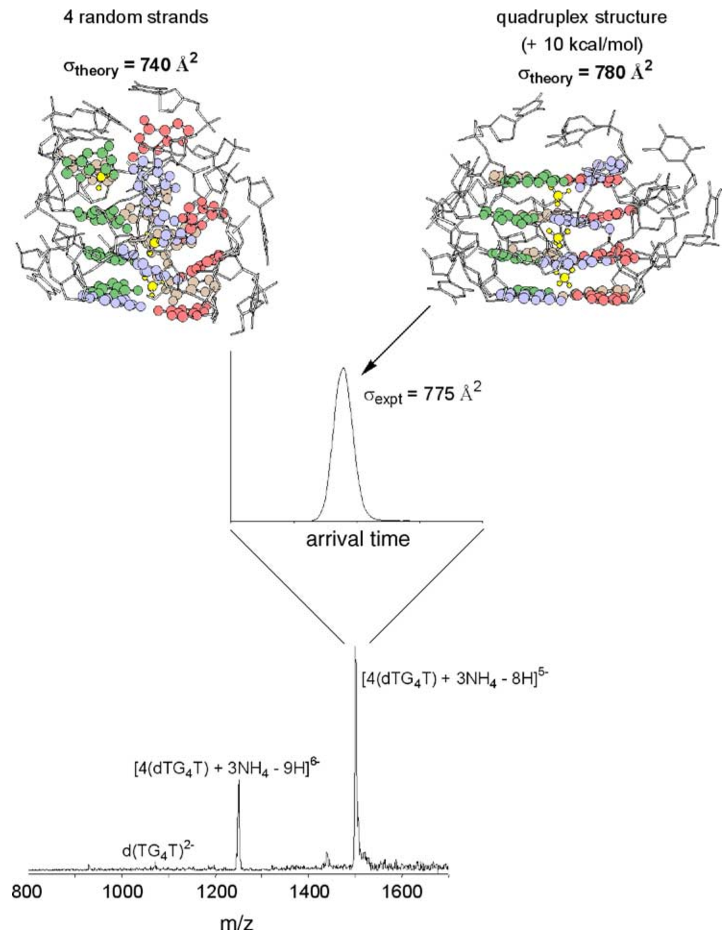

Fig. 11. ESI mass spectrum of d(TGGGGT). Also shown are the ATD and theoretical structures for $[4(dTGGGGT) + 3NH_4 - 8H]^{5-}$. The globular structure (with the four oligonucleotide strands randomly placed together) is the lowest in energy but the cross-section of the quadruplex structure matches best with experiment.

Fig. 11 shows the mass spectrum, ATD, and theoretical structures of a d(TGGGGT) quadruplex obtained in our lab. The four-stranded species can be readily identified by its *m/z* value and is the most intense peak in the mass spectrum. The quadruplex is only observed in the presence of NH_4^+ and previous mass spectrometry studies have observed similar results [61]. The ATD for the 5^- charge state quadruplex is also shown in Fig. 10, yielding a single, symmetric peak and an experimental cross-section of 775 ± 15 Å^2.

In the theoretical modeling, the $[4(dTGGGGT) + 3NH_4 - 8H]^{5-}$ ion was initially placed in one of two starting geometries. If the four hexanucleotide strands are randomly placed together, the resulting lowest energy structure from the simulated annealing is shown in the upper left corner of Fig. 11. This final structure is not a “quadruplex” (which consists of a planar array of G-quartets) but the guanine bases in each strand do preferentially stack on top of each other. However, the cross-section of this structure is 5% smaller than experiment (a value that falls outside acceptable error limits). If the four hexanucleotide strands are initially placed in a “quadruplex” geometry (using published X-ray data [62]), the resulting lowest energy structure from the simulated annealing is shown in the upper right corner of Fig. 11. The final structure remains in a “quadruplex” geometry with the three NH_4^+ groups situated between each G-quartet. This structure is ∼10 kcal/mol higher in energy than the random form, but its cross-section agrees much better with experiment.

In the duplex studies, theory predicted that the lowest energy solvent-free structures were globular but the ion mobility experiments indicated helical duplex structures were the sole species observed for larger strands. The reasoning behind this result was that the duplexes were initially helical in solution and retained that structure for a limited time in the gas phase. A similar situation may be occurring here. Although the quadruplex structure is not lowest in energy in the gas-phase, it may be the preferred structure in solution and is retaining that structure in the gas phase. What is interesting though, is the fact that the ion mobility results clearly indicate that the four-stranded d(TGGGGT) system must be a quadruplex, as four randomly placed strands (or non-specific tetramers), will not yield a structure that matches experimental data.

4. Conclusions

Although the results presented in this article are far from comprehensive and much work remains to be done, they do provide several important insights into the gas-phase conformations of DNA duplexes and higher-order complexes and some of the factors that affect these conformations:

(1) Watson–Crick pairing is enhanced in dinucleotides that are cationized by d^{10} metals.
(2) Duplexes can retain helical structures on the ms time scale in the gas phase. These helical structures are size dependent (first appearing at the 8-mer length), base dependent (with A·T pairs preferentially broken over C·G pairs), and sequence dependent (with A·T pairs on the ends of the duplexes more easily disrupted than if they are in the middle of the duplex).
(3) Quadruplex structures are observed for G-rich oligonucleotides in the presence of ammonium ions.
(4) In almost all cases, the structures observed in the ion mobility experiments are not the lowest energy structures predicted by AMBER 7 calculations.
(5) Thus, solution phase structures are retained in the absence of solvent with relatively minor intramolecular changes in conformation.

Acknowledgement

The support of the National Science Foundation under grant CHE0140215 is gratefully acknowledged.

References

[1] W. Saenger, Principles of Nucleoc Acid Structure, Springer, New York, 1988.
[2] R.R. Sinden, DNA Structure and Function, Academic Press, San Diego, 1994.
[3] B. Ganem, Y.T. Li, J.D. Henion, Tetrahedron Lett. 34 (1993) 1445.
[4] K.J. Light-Wahl, D.L. Springer, B.E. Winger, C.G. Edmonds, D.G. Camp, B.D. Thrall, R.D. Smith, J. Am. Chem. Soc. 115 (1993) 803.
[5] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.M. Whitehouse, Science 246 (1989) 64.
[6] M.J. Doktycz, S. Habibigoudarzi, S.A. McLuckey, Anal. Chem. 66 (1994) 3416.
[7] P.A. Limbach, P.F. Crain, J.A. McCloskey, Curr. Opin. Biotechnol. 6 (1995) 96.
[8] J.A. Loo, Mass Spectrom. Rev. 16 (1997) 1.
[9] B. Guo, Anal. Chem. 71 (1999) 333R.
[10] S.A. Hofstadler, R.H. Griffey, Chem. Rev. 101 (2001) 377.
[11] V. Gabelica, E. DePauw, J. Mass Spectrom. 36 (2001) 397.
[12] V. Gabelica, E. DePauw, Int. J. Mass Spectrom. 219 (2002) 151.
[13] P.D. Schnier, J.S. Klassen, E.F. Strittmatter, E.R. Williams, J. Am. Chem. Soc. 120 (1998) 9605.
[14] R.H. Griffey, M.J. Grieg, H. An, H. Sasmor, S. Manalili, J. Am. Chem. Soc. 121 (1999) 474.
[15] E.A. Mason, E.W. McDaniel, Transport Properties of Ions in Gases, Wiley, New York, 1988.
[16] T. Wyttenbach, P.R. Kemper, M.T. Bowers, Int. J. Mass Spectrom. 212 (2001) 13.
[17] E. Shammel-Baker, J. Gidden, D.P. Fee, P.R. Kemper, S.E. Anderson, M.T. Bowers, Int. J. Mass Spectrom. 227 (2003) 205.
[18] J. Gidden, P.R. Kemper, E. Shammel, D.P. Fee, S. Anderson, M.T. Bowers, Int. J. Mass Spectrom. 222 (2003) 63.
[19] P.R. Kemper, M.T. Bowers, J. Am. Soc. Mass Spectrom. 1 (1990) 197.
[20] F. Hillenkamp, M. Karas, R.C. Beavis, B.T. Chait, Anal. Chem. 63 (1991) 1193A.
[21] D.A. Case, D.A. Pearlman, J.W. Caldwell, T.E. Cheatham III, J. Wang, W.S. Ross, C.L. Simmerling, T.A. Darden, K.M. Merz, R.V. Stanton, A.L. Cheng, J.J. Vincent, M. Crowley, V. Tsui, H. Gohlke, R.J. Radmer, Y. Duan, J. Pitera, I. Massova, G.L. Seibel, U.C. Singh, P.K. Weiner, P.A. Kollman, AMBER 7, University of California, San Francisco, 2002.
[22] P.J. Stevens, F.J. Devlin, C.F. Chablowski, M.J. Frisch, J. Phys. Chem. 98 (1994) 11623.
[23] A.K. Becke, J. Chem. Phys. 98 (1993) 5648.
[24] T. Wyttenbach, G. von Helden, J.J. Batka, D. Carlat, M.T. Bowers, J. Am. Soc. Mass Spectrom. 8 (1997) 275.
[25] T. Wyttenbach, M. Witt, M.T. Bowers, J. Am. Chem. Soc. 122 (2000) 3458.
[26] A.A. Shvartsburg, M.F. Jarrold, Chem. Phys. Lett. 261 (1996) 86.
[27] J. Gidden, T. Wyttenbach, J.J. Batka, P. Weis, A.T. Jackson, J.H. Scrivens, M.T. Bowers, J. Am. Soc. Mass Spectrom. 10 (1999) 883.
[28] J. Gidden, M.T. Bowers, J. Am. Soc. Mass Spectrom. 14 (2003) 161.
[29] J. Gidden, M.T. Bowers, J. Phys. Chem. B 107 (2003) 12829.
[30] J. Gidden, A. Ferzoco, E. Shammel Baker, M.T. Bowers, J. Am. Chem. Soc., in press.
[31] J. Gidden, M.T. Bowers, Eur. Phys. J. D 20 (2002) 409.
[32] E. Shammel Baker, J. Gidden, A. Ferzoco, M.T. Bowers, Phys. Chem. Chem. Phys. 6 (2004) 2786.
[33] E. Hammarsten, Biochem. Z 144 (1924) 383.
[34] B. Rosenberg, L. VanCamp, J.E. Trosko, V.H. Mansour, Nature 222 (1969) 385.
[35] H.C. Harder, B. Rosenberg, Int. J. Cancer 6 (1970) 207.
[36] B. Lippert, Prog. Inorg. Chem. 37 (1989) 1.
[37] J. Reedijk, Chem. Rev. 99 (1999) 2499.
[38] J.R. Williamson, Cell 23 (1994) 703.
[39] N. Spackova, I. Berger, J. Sponer, J. Am. Chem. Soc. 121 (1999) 5519.
[40] G.N. Parkinson, M.P.H. Lee, S. Neidle, Nature 417 (2002) 876.
[41] J.S. Ren, X.G. Qu, J.O. Trent, J.B. Chaires, Nucleic Acids Res. 30 (2002) 2307.
[42] B. Lippert, Coord. Chem. Rev. 200–202 (2002) 487.
[43] J.V. Burda, J. Sponer, P. Hobza, J. Phys. Chem. 100 (1996) 7250.
[44] J.V. Burda, J. Sponer, J. Leszczynski, P. Hobza, J. Phys. Chem. B 101 (1997) 9670.
[45] E. Shammel Baker, J. Gidden, M.J. Manard, M.T. Bowers, in preparation.
[46] R.G. Pearson, J. Am. Chem. Soc. 85 (1963) 3533.
[47] G. Klopman, J. Am. Chem. Soc. 90 (1968) 223.
[48] R. Langridge, H.R. Wilson, C.W. Hooper, M.H.F. Wilkins, L.D. Hamilton, J. Mol. Biol. 2 (1960) 19.
[49] W. Fuller, M.H.F. Wilkins, H.R. Wilson, L.D. Hamilton, S. Arnott, J. Mol. Biol. 12 (1965) 60.
[50] A.H. Wang, G.J. Quigley, F.J. Kolpak, J.L. Crawford, J.H. van Boom, G. Van der Marel, A. Rich, Nature 282 (1979) 680.
[51] The mass spectra of each duplex showed duplex peaks with charge states equal to half its length (i.e., the 14-mer had a charge state of 7^-). Thus, the 6, 10, 14, and 18-mer duplexes had odd charge states and they could be clearly identified in the mass spectra based on their m/z values. The 8-mer duplex had a charge state of 4^-, which has the same m/z value as a single stranded 8-mer with a charge state of 2^-. In the d(CG) series, the 8-mer duplex could be identified based on the $\Delta m/z$ of the sodiated adducts (see ref. [47]).

The same sodium splitting was not observed in the d(AT) 8-mer mass spectra.
[52] M. Rueda, S.G. Kalko, F.J. Luque, M. Orozco, J. Am. Chem. Soc. 125 (2003) 8007.
[53] H. Htun, J.E. Dahlberg, Science 241 (1988) 1791.
[54] H. Htun, J.E. Dahlberg, Science 243 (1989) 1571.
[55] J.N. Glover, C.S. Farah, D.E. Pulleyblank, Biochemistry 29 (1990) 11110.
[56] D. Sen, W. Gilbert, Nature 334 (1988) 364.
[57] K. Fukushima, H. Iwahashi, Chem. Commun. (2000) 895.
[58] I. Manet, L. Francini, S. Masiero, S. Pieraccini, G.P. Spada, G. Gottarelli, Helv. Chim. Acta 84 (2001) 2096.
[59] T. Aggerholm, S.C. Nanita, K.J. Koch, R.G. Cooks, J. Mass Spectrom. 38 (2003) 87.
[60] D.R. Goodlett, D.G. Camp II, C.C. Hardin, M. Corregan, R.D. Smith, Biol. Mass Spectrom. 22 (1993) 181.
[61] F. Rosu, V. Gabelica, C. Houssier, P. Colson, E. De Pauw, Rapid Commun. Mass Spectrom. 16 (2002) 1729.
[62] G.R. Clark, P.D. Pytel, C.J. Squire, S. Neidle, J. Am. Chem. Soc. 125 (2003) 4066.