Eur. Phys. J. D **20**, 409–419 (2002)
DOI: 10.1140/epjd/e2002-00170-7

THE EUROPEAN PHYSICAL JOURNAL D

Gas-phase conformational and energetic properties of deprotonated dinucleotides

J. Gidden and M.T. Bowers[a]

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

Received 17 May 2002
Published online 13 September 2002 –

Abstract. Ion mobility experiments and molecular modeling calculations were used to investigate the gas-phase conformations and folding energetics of 16 deprotonated dinucleotides. $[M–H]^-$ ions were formed by MALDI and their collision cross-sections measured in helium using ion mobility based techniques. Cross-sections of theoretical structures, generated by molecular mechanics/dynamics calculations, were compared to the experimental values for conformational identification of the dinucleotides. Temperature dependent measurements and kinetic theory were also used to obtain energetic and dynamic data concerning the folding properties of the dinucleotides. Three distinct families of conformations, with significantly different collision cross-sections, were identified: a "stacked" family in which the two nucleobases stack; an "H-bonded" family in which the two nucleobases stay in the same plane and are hydrogen-bonded to each other; and an "open" family in which the two nucleobases are separated from each other. At temperatures $\geq$ 300 K these conformers rapidly interconvert in most systems, but they can be separated and individually observed in the lower temperature (80–200 K) experiments. The types and relative amounts of each conformer observed, and the temperature at which they can be separated, are base and sequence dependent. Theoretical modeling of the temperature-dependent data was used to determine isomerization barrier heights between the various conformers and yielded values between 0.8–12.9 kcal/mol, depending on the dinucleotide.

PACS. 07.75.+h Mass spectrometers – 87.15.Cc Folding and sequence analysis – 87.15.He Dynamics and conformational changes

1 Introduction

Investigations into the structures and conformations of DNA are essential towards understanding its fundamental physical and chemical properties. For example, DNA replication mechanisms, DNA–protein binding processes, and DNA–drug interactions are all strongly dependent on the conformational properties of the DNA molecule (as well as the protein and drug) [1,2]. Most structural analyses of DNA involve condensed-phase measurements, such as NMR (*in vitro*) [3–5] or X-ray crystallography (in crystalline solids) [6], which have provided an abundant amount of information concerning the conformational properties of oligonucleotides. Less attention, however, has been paid to the gas-phase structural properties of oligonucleotides. One reason is that, until recently, oligonucleotides could not easily be transferred, intact, into the gas phase. Another reason is that DNA is normally found in the cellular environment and so condensed-phase experiments are usually considered appropriate choices for analysis.

However, gas-phase studies of oligonucleotides should not be easily disregarded. High-resolution crystal data is needed to distinguish between small conformational changes and important structural information can often be hidden in the X-ray data. NMR results are often solvent dependent and typically give only averaged conformational data. Intrinsic structural and energetic data about the oligonucleotide itself, such as base–base interactions, can only come from solvent-free environments. Additionally, more theoretical modeling is accompanying experimental studies due to advances in computing capabilities. Unlike solution-phase or solid-state experiments, gas-phase structural data are ideal for direct comparison to theoretical structures.

Continuing improvements in matrix-assisted laser desorption ionization (MALDI) [7] and electrospray ionization (ESI) [8] sources have helped mass spectrometry emerge as an effective tool for characterizing biological molecules in the gas phase. MS data on DNA containing 2–100 bases have been reported [9–11]. Most of these studies

[a] e-mail: bowers@chem.ucsb.edu

have focused on obtaining sequence information, utilizing the speed of the technique and its ability to analyze small (picomole) sample sizes. Sequencing data is often obtained by fragmenting the oligonucleotide ion *via* collision-induced dissociation (CID) or post source decay [12–14]. The resulting fragment ions can yield extensive information about the structure and sequence of the parent ion. In fact, libraries of CID data of peptide and protein ions have been collected to aid in the identification of unknown samples. However, it has also been shown that the higher order structures of the peptides and proteins can significantly affect their CID spectra [15,16] and hence influence structural and sequence assignments. While there are numerous studies that have probed the gas-phase conformations of peptides and proteins, almost no data is available on the gas-phase conformations of oligonucleotides (whose sequence assignments may also depend on conformation).

In this paper, gas-phase conformational and energetic data of 16 deprotonated dinucleotides will be presented. (Since so little is known about the gas-phase conformations of oligonucleotides [17], the simplest systems – dinucleotides – were chosen.) The necessary information was obtained through ion mobility measurements [18,19] and molecular modeling calculations. Ion mobility is based on the measurement of the amount of time it takes for a pulse of mass-selected ions to drift through a buffer gas under the influence of a weak, uniform electric field. This drift time is dependent on the ions' collision cross-sections, and hence their geometric shape. Ions that are tightly folded have smaller collision cross-sections and will drift faster than ions that are more extended and have larger collision cross-sections. If the ions have multiple conformations, with significantly different collision cross-sections, each conformer may be separated from the others as it drifts through the buffer gas/electric field [20–23]. Therefore, these ion mobility measurements are sometimes dubbed "ion chromatography". Various computational methods are then used to generate trial structures of the ions and calculate their corresponding cross-sections for comparison to experiment and conformational identification of the ions. Ion mobility methods have been successfully used to analyze the conformations of a number of different peptides [24–29], proteins [22,30,31], and other biological polymers [32–34]. A brief communication of our ion mobility results on the conformational and energetic analysis of two dinucleotides, dTG^- and dGT^-, has been published [35]. In this paper, we extend those methods to the analyses of all 16 dinucleotides.

2 Experiment

Detailed descriptions of the MALDI ion source [36] and the instrumentation used for the mobility measurements [37] have been published so only a brief description will be given here. Desalted dinucleotides were purchased from Sigma–Genosys and used without further purification. A schematic of the dinucleotides along with the four nucleobases is shown in Figure 1. 2,5-dihydroxybenzoic acid (DHB) was used as the matrix and methanol as the solvent. The matrix solution was prepared at a concentration of 100 mg/ml and the dinucleotide solution at a concentration of $\sim$ 1 mg/ml. Approximately 100 μl of DHB and 100 μl of the dinucleotide were applied to the sample target (a stainless steel cylinder 1.25 cm long and 1.25 cm in diameter) and dried.

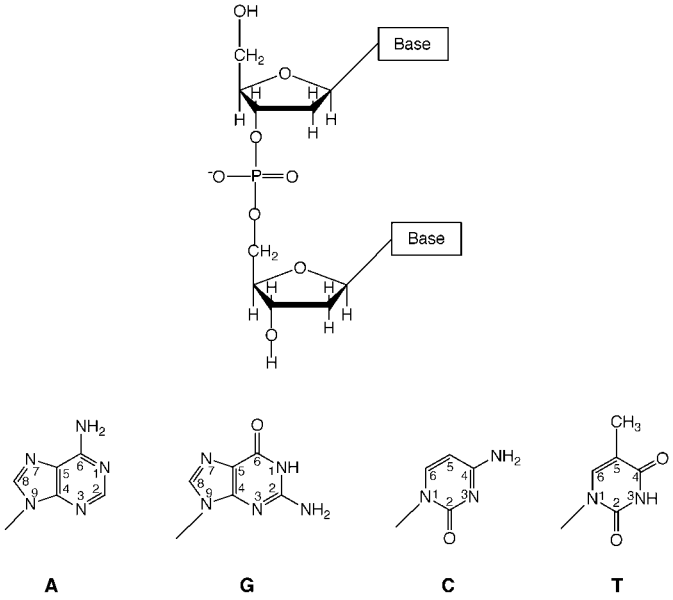

Fig. 1. Schematic diagram of the deprotonated dinucleotide and the four nucleobases.

A nitrogen laser (λ = 337 nm) operated at 100 Hz with a pulse width of 10–20 ns and pulse energy of 1–2 mJ was used to generate the ions in the MALDI source. Ions exiting the source are accelerated to 5 kV and mass analyzed with a reverse-geometry, double-focusing mass spectrometer similar to a V.G. ZAB-2F [38]. The appropriate $[M–H]^-$ ions are mass selected, decelerated, and injected at low energies into a 4 cm long, copper drift cell [37] filled with $\sim$ 3 torr of He. The ions drift through the cell under the influence of a weak electric field and are collected as a function of time, yielding an arrival time distribution or ATD. The laser pulse triggers the "timer" and data is collected on a multi-channel scalar with a 2-μs channel width. The temperature of the drift cell can be varied from 80 to 580 K by flowing warmed or cooled nitrogen through passages surrounding the cell.

The mobility, K_0, of the ion is determined from a series of ATDs measured at different electric field strengths (5–25 V/cm) using

$$t_A = \left[l^2 \frac{273}{760T} \left(\frac{1}{K_0} \frac{p}{V} \right) \right] + t_0 \qquad (1)$$

where t_A is the arrival time (taken from the center of the ATD peak), l is the length of the drift cell, T is temperature, p is the pressure of the He gas, V is the electric field strength, and t_0 is the time the ions spend outside of the drift cell [39]. A plot of t_A *vs.* p/V yields a straight line with a slope proportional to K_0 and an intercept of t_0. From the mobility, the ion's collision cross-section, $\Omega^{(1,1)}$,

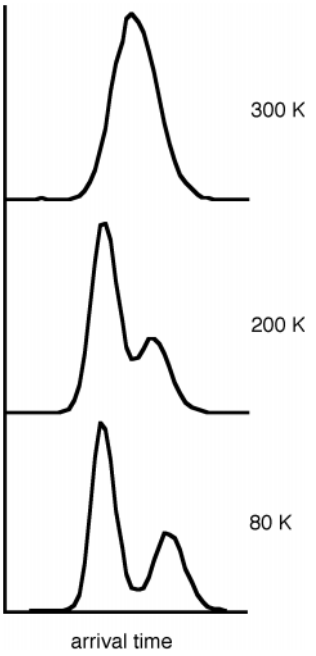

Fig. 2. Arrival time distributions for dCG measured at different temperatures. Multiple peaks signify the presence of multiple conformations. As the temperature increases, the conformers begin to interconvert as they drift through the mobility cell and a single peak appears in the ATD.

can be obtained using kinetic theory [39]

$$\Omega^{(1,1)} = \frac{3q}{16N_0}\left(\frac{2\pi}{\mu kT}\right)^{1/2}\frac{1}{K_0} \qquad (2)$$

where q is the ion charge, N_0 is the number density of He at STP, μ is the ion–He reduced mass, k is Boltzmann's constant, and T is temperature. Hence, the ion's arrival time is inversely proportional to its mobility and directly proportional to its collision cross-section with He. Compact ions with small cross-sections will, therefore, have shorter arrival times than more extended ions with larger collision cross-sections. If the ions have multiple conformations with significantly different cross-sections that do not rapidly interconvert, each conformer will drift through the cell with different mobilities and appear as different peaks in the ATDs [20,21,40]. Under our experimental conditions, ions with cross-sections differing by 3% or more can be resolved.

Conformational identification of the ions is achieved by comparing the cross-sections of theoretical structures with those obtained from the ATDs. The AMBER 6.0 molecular mechanics/dynamics package [41] was used to generate model structures for the 16 dinucleotide ions. A series of simulated annealing and energy minimization cycles yielded 100 low energy structures for each dinucleotide. The process is as follows. An initial structure is energy minimized, run through a 30-ps MD simulation at 800 K, cooled to 0 K through another 10-ps MD simulation, and energy minimized again. This final structure is saved and used as the starting structure for another annealing/minimization cycle. The entire process is repeated until 100 structures are obtained. The angle-averaged collision cross-section of each structure was calculated with a previously developed temperature-dependent projection model [20,36,42] that has yielded reliable cross-sections for a number of biological [24–26] and synthetic [36,40,43,44] polymers. Scatter plots of cross-section *vs.* energy are then used to help identify the ion conformation(s) observed in the ATDs.

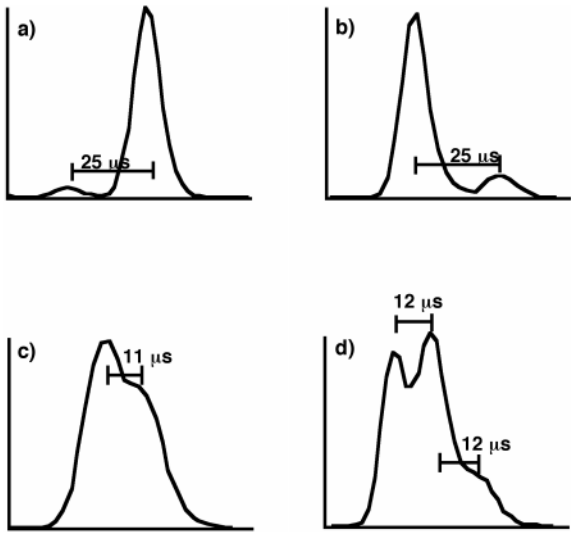

Fig. 3. 80 K ATDs for (a) dAT, (b) dTA, (c) dGA, and (d) dGC. All 16 dinucleotides have 80 K ATDs that resemble one of the four shown here (see Tab. 1 and text).

3 Results/discussion

3.1 Conformational properties of the dinucleotides

Typical arrival time distributions (ATDs) obtained for the dinucleotides are shown in Figure 2 using dCG as an example. These ATDs were measured under similar experimental conditions except for the temperature of the cell. At high temperatures (300 K and above), only one peak is observed in the ATDs, but as the temperature is lowered to 80 K, multiple peaks begin to appear in the spectra. The number of peaks present in each 80 K ATD and their relative intensities are system dependent, but the dinucleotides can be grouped into four distinct categories based on similarities in their 80 K ATDs. These groups are shown in Figure 3 using dAT, dTA, dGA, and dGC as examples.

In the first group (Fig. 3a), two peaks are present in the 80 K ATD that are separated by ~ 25 μs. The shorter-time peak is always *less intense* than the longer-time peak in this group, but the ratios of the peaks vary from 1:80 to 1:6. dAT, dCT, dGT, dTT, and dTG are part of this group. The second group (Fig. 3b) also has two peaks in the 80 K ATDs that are separated by ~ 25 μs. However, the shorter-time peak is always *more intense* than

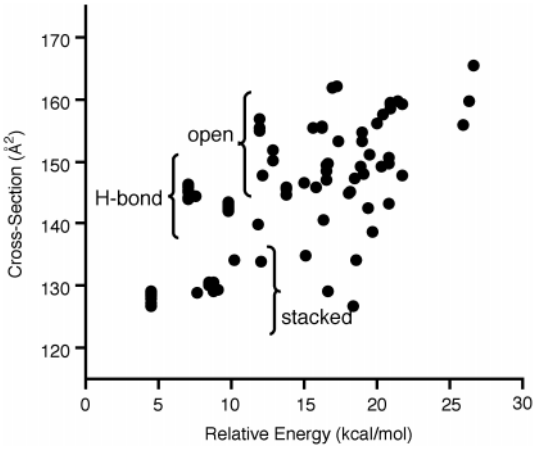

Fig. 4. Scatter plot of 300 K cross-section *vs.* energy for the 100 structures calculated for dCC using the simulated annealing procedure. Each point represents 1 theoretical structure. Three major types of conformations are identified on the plot. See Figure 5 for examples.

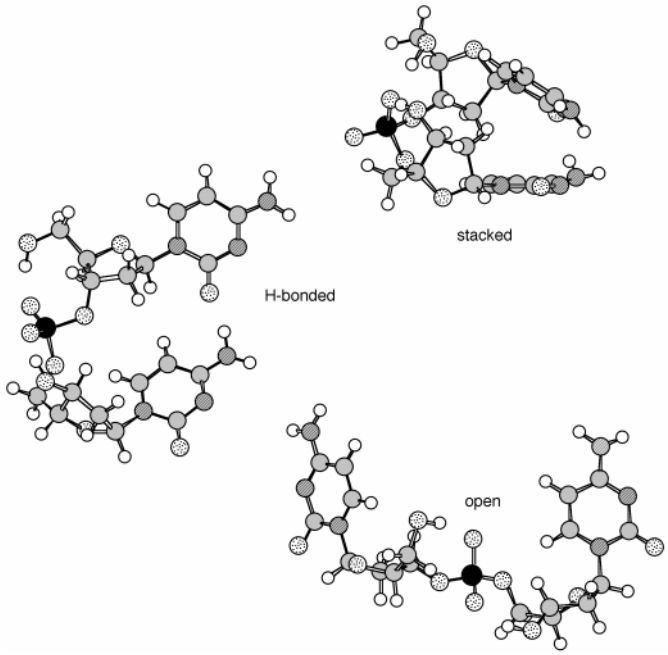

Fig. 5. Representatives of the three families of conformations found in the scatter plots. Carbon atoms are gray, oxygens are spotted, nitrogens are striped, hydrogens are white, and the phosphorous atom is black. The "stacked" conformer has the smallest cross-section while the "open" conformer has the largest cross-section.

the longer-time peak in this group, with peak ratios varying from 18:1 to 5:2. dAA, dCA, dTA, dAG, and dCG fall into this group.

In the third group (Fig. 3c), two peaks are observed in the 80 K ATDs that are separated by only ~ 12 μs. dGA and dGG are the only dinucleotides in this group and have peak ratios of 4:3 and 1:3, respectively. The fourth group (Fig. 3d) has *three* peaks present in the 80 K ATDs, each of which are separated by ~ 2 μs. The middle-time peak is the most intense peak in this group. dAC, dCC, dGC, and dTC are members of this group and have peak ratios of 9:10:1, 1:2:1, 7:8:1, and 1:8:2, respectively.

Multiple peaks in ATDs indicate that the dinucleotides have multiple conformations with significantly different collision cross-sections that do not rapidly interconvert on the experimental time scale (500 ± 200 μs) [35,40]. For the ATDs shown in Figure 3, a time difference of ~ 25 μs corresponds to a cross-section difference of 20–22 Å^2 while a time difference of ~ 12 μs corresponds to a 10–12 Å^2 difference in cross-section. In order to determine the conformational family of each peak in the ATD, extensive theoretical modeling of the dinucleotides was performed. The energies and cross-sections of 100 structures of each dinucleotide were calculated using the procedures described in the "Experimental" section. A "scatter plot" of cross-section *vs.* energy of the structures is then used to help identify the conformations observed in the ion mobility experiments. The scatter plot for dCC, in which cross-sections were calculated using a temperature of 300 K, is shown in Figure 4.

Each point in the plot represents one theoretical structure. Minor variations in structure and the statistical nature of the cross-section calculations produce some of the "scatter" in the plot, but three distinct families of conformers, with different cross-sections, can be identified: stacked, H-bonded, and open. Examples of each family are shown in Figure 5. The "stacked" family, in which the two nucleobases are stacked, have the smallest cross-section and are usually predicted to be lowest in energy. The bases generally stack so that the carbonyl oxygen or ring nitrogen on one base is situated over the NH_2 group (or NH on thymine) on the other base. This is similar to the types of overlap observed for free bases [45]. The hydrogen atoms on the amino group can bend out of plane to interact with the =O or –N– atom on the other base and hydrogen bonds between the bases are observed for most of the "stacked" dinucleotides. (H-bonds are not observed in the "stacked" forms of dAA and dGA where the NH_2 group remained planar and was situated over the aromatic system on the other base.) However, the ion mobility results indicate that these hydrogen bonds in the "stacked" form are not the primary factor in determining whether a dinucleotide prefers to stack (this will be discussed later).

In the "H-bonded" family, the two bases are roughly in the same plane and are hydrogen-bonded to each other through an amino hydrogen on one base and a carbonyl oxygen (C, G, or T) or ring nitrogen (N1 or N3 in A) on the other base. These conformers are similar in energy to the "stacked" family (within 2 kcal/mol) but have 10–12 Å^2 larger cross-sections. In the "open" family, the two bases are completely separated from each other. These conformers have the largest cross-sections, 20–25 Å^2 larger than the "stacked" conformers, and are predicted to be slightly higher in energy (1–5 kcal/mol) than the "stacked" forms.

Most of the dinucleotides have similar scatter plots and conformations as those shown in Figures 4 and 5, but there are a few exceptions worth noting. First, in the

purine–purine systems (dAA, dAG, dGA, and dGG) the bases in the "stacked" conformer are more parallel to each other than that shown in Figure 5 for dCC. In the purine–pyrimidine and pyrimidine–pyrimidine systems, the bases are actually angled towards one another in the "stacked" form (as seen in Fig. 5). Second, the scatter plots for dGA and dGG indicate that the "open" conformers for these two dinucleotides are ~ 10 kcal/mol higher in energy than the "stacked" or "H-bonded" conformers (which have similar energies). This is over twice the difference calculated for the other 14 dinucleotides. Finally, no "H-bonded" conformations are theoretically predicted for dAT, dCT, or dTT. The hydrogen bonds between the two bases in the "H-bonded" conformers usually involve the carbonyl oxygen (or ring nitrogen in A) on the 5′ base and an amino hydrogen on the 3′ base. Thymine (T) is the only base without an NH_2 group and so when it is in the 3′ position, hydrogen bonds between it and the 5′ base are not structurally accessible. The CH_3 group on the 3′T is usually pointed toward the 5′ base. However, if the 5′ base is guanine (G) then the hydrogen bonding between bases involves the NH_2 group on guanine and the carbonyl oxygen or ring nitrogen (A) on the 3′ base. Therefore, stable "H-bonded" conformations are predicted for dGT but not dAT, dCT, nor dTT.

Based on the scatter plot data and the resulting families of conformations, the peaks in the 80 K ATDs (Fig. 3) can now be identified. Identification is usually made by direct comparison of the calculated cross-sections of the theoretical structures with the experimental values determined from the ATDs. Unfortunately, absolute cross-section comparisons are difficult in this case. The cross-section calculations require the specification of a temperature due to the ion–He interaction potential. Since this potential is relatively shallow it has little effect on the cross-section at temperatures of 300 K or higher. However, when the temperature falls below 300 K the interaction potential well starts to become comparable to kT and "increases" in experimental cross-sections of 50% or more have been observed near 80 K (see Ref. [42]). In calculating cross-sections at these low temperatures, the effect of the attractive part of the potential surface is amplified. Since we do not have accurate multidimensional surfaces, uncertainty in the calculated cross-sections is also amplified. Hence, quantitative, absolute cross-section comparisons with experiment are less useful but relative cross-section differences between conformers are still manageable and will form the basis of our assignments.

The fact that the 80 K ATDs show multiple peaks while, for most dinucleotides, the 300 K ATDs show single peaks indicates that multiple conformers exist but are rapidly isomerizing at 300 K (this will be discussed in more detail in the next section). Thus, the experimental cross-sections abstracted from the 300 K ATDs are weighted averages of the cross-sections of the individual conformers (see Sect. 3.3 and Fig. 12). While theory can give unambiguous predictions of the cross-sections of individual conformers at 300 K, the ion mobility experiments cannot clearly distinguish between them at this temperature. However, in the case of dCC, which will be discussed in more detail later, an unambiguous comparison between experiment and theory can be made because the stacked conformer does not isomerize to the H-bonded or open conformers at 300 K. From the scatter plot in Figure 4, a theoretical cross-section of 130 ± 2 Å^2 is predicted for the stacked conformer while the experimental cross-section from the 300 K ATD is 132 Å^2. Clearly, the agreement is excellent and, from past experience, there is every reason to believe that similar agreement will hold true for all dinucleotides and conformers (see Refs. [24–26,36,40,43,44]).

Table 1. Categories the dinucleotides can be grouped into based on the conformations observed in their 80 K ATDs.

Observed conformers	dinucleotides
stacked/open[a]	dAT, dCT, dGT, dTT, dTG
stacked[a]/open	dAA, dCA, dTA, dAG, dCG
stacked[b]/H-bonded[b]	dGA, dGG
stacked/H-bonded[a]/open	dAC, dCC, dGC, dTC

[a]Dominant peak in the ATD, [b]the "stacked" peak is slightly dominant for dGA while the "H-bonded" peak is dominant for dGG.

Table 1 summarizes the 80 K ATD peak identification for all of the dinucleotides based on comparison of cross-sectional differences observed from experiment and theory. In Figures 3a and 3b, the shortest-time peak in the ATD can be assigned to the "stacked" conformer (smaller cross-section, shorter arrival time) while the longest-time peak is consistent with the "open" conformer. There is no indication that significant (if any) amounts of the "H-bonded" conformers are present for these dinucleotides (see Tab. 1 for list). The time difference between the two peaks yields a cross-section difference of 20–22 Å^2, a magnitude that agrees very well with the theoretical predictions for the "stacked" and "open" conformers. As will be shown later, if significant amounts of the "H-bonded" conformers were present in these dinucleotides, but not separated from the "stacked" or "open" conformers at 80 K, the two peaks in the ATDs would be closer together.

In Figure 3c, the two peaks in the ATD are consistent with the "stacked" (shorter-time) and "H-bonded" (longer-time) conformers. In this case, the "open" conformer is not observed. In Figure 3d, all three families of conformers are observed in the 80 K ATD. The shortest-time peak can be assigned to the "stacked" conformer, the middle peak is the "H-bonded" conformer, and the longest-time peak is the "open" conformer. This last group is the only one in which all three families of conformers are observed in the 80 K ATDs.

Two general trends are apparent in the data presented in Table 1. The first trend concerns the types of conformers observed in each 80 K ATD. Even though theory predicts three distinct, low-energy conformations for 13 of

the 16 dinucleotides (dAT, dCT, and dGT are the exceptions), all three conformers are experimentally observed for only 4 of the dinucleotides. The "stacked" form is the only conformer that is observed in the 80 K ATDs of all 16 dinucleotides (although the amount for dCT and dTT is extremely small). The "open" conformer is present in the ATDs of 14 of the 16 dinucleotides, with dGA and dGG as the lone exceptions. The "H-bonded" conformer, however, is experimentally observed for only 6 of the dinucleotides.

It is somewhat surprising that the "stacked" conformer is observed for all of the dinucleotides, while the "H-bonded" forms are observed for only a few of them. Studies on free base pairs and mononucleotide dimers have shown that hydrogen bonding between the bases is more favorable in the gas-phase while base stacking is favored only in solution [46–50]. Of course, as isolated base pairs, the two nucleobases are free to orient themselves as they wish to maximize the interaction between them. In the dinucleotide case, the bases are anchored to the phosphate–sugar backbone and have a more limited range of movement. For example, Watson-Crick hydrogen bonding, which is prevalent for A·T and C·G base pairs, is not possible in these dinucleotides because the bases cannot position themselves to allow this type of interaction to occur. Additionally, H-bonding in the free bases typically involves multiple sites. In the dinucleotides, only one H-bond between the two bases is usually observed. Consequently, stacking interactions between the bases, instead of H-bonding, appears to be preferable for these dinucleotides.

As mentioned previously, dGA and dGG are the only dinucleotides in which the "open" conformer is not experimentally observed in the 80 K ATDs. dGA and dGG also are the only dinucleotides in which theory predicts that the "open" conformer is significantly higher in energy ($\sim$ 10 kcal/mol) than the "stacked" or "H-bonded" forms. This result is most likely due to an increase in interaction between the two purine bases rather than a decreased stability of the "open" form. Theoretical studies on isolated base pairs have shown that base stacking tends to favor purine·purine > purine·pyrimidine > pyrimidine·pyrimidine systems [2,51]. It is interesting to note that the ATDs for the other two purine–purine systems, dAA and dAG, show that these dinucleotides will "open up", although the stacked form is still favored for these systems.

One reason for the different behavior between the four purine–purine systems may be due to differences in the number and types of hydrogen bonds between the bases. In dGG, there are actually two H-bonds between the guanine bases, one between the NH_2 group on the 5′G and the carbonyl oxygen on the 3′G and another between the NH_2 group on the 5′G and N7 on the 3′G. dGG was the only dinucleotide in which two H-bonds between the bases were observed in the "H-bonded" family. (G is also a very polar base and H-bonded G·G base pairs are almost as stable as G·C Watson-Crick base pairs [49].) For dGA, the NH_2 on the 5′G hydrogen bonds to N1 on the 3′A and to the hydroxyl oxygen on the 3′ end of the sugar (dGC also had this latter type of hydrogen bonding). In dAG and dAA, only one hydrogen bond involving the bases is observed. While H-bonds are not necessarily strong bonds, the addition of just one extra H-bond may be enough to keep dGG and dGA from "opening up" in the time they spend in the drift cell (especially at low temperatures where the internal energy in the system is low).

The second trend apparent in the data in Table 1 is that the conformational preferences of the dinucleotides are strongly dependent on the identity of the 3′ base. Dinucleotides with a 3′T (dAT, dCT, dGT, dTT), for example, favor "open" conformations. Dinucleotides with a 3′A prefer "stacked" conformations and those with a 3′C prefer "H-bonded" conformations (and are the only ones in which all three conformers are experimentally observed at 80 K). Conformational preferences of dinucleotides with a 3′G tend to be more dependent on the 5′ base. dAG and dCG prefer "stacked" conformations, dGG prefers the "H-bonded" conformation, and dTG prefers the "open" conformation.

The exact reason why the 3′ base directs the conformational preferences of the dinucleotides is not clear at this time. The AMBER calculations do not show any major structural or relative energetic variations in the "stacked" or "open" families that are strongly dependent on the 3′ base. (The atoms involved in the hydrogen bonding in the "H-bonded" conformers always involve the NH_2 atoms on the 3′ base, unless G is the 5′ base.) For example, in the "stacked" conformations of dTA and dAT, the relative orientations of the A and T bases are the same in both dinucleotides. The O2 atom on T is situated over the NH_2 group on A and the N1 atom on A is situated over the NH group on T. The O2$\cdots$H and N1$\cdots$H distances are also similar in dTA and dAT and the bases are angled towards one another in a similar manner for both dinucleotides. Yet, the experiments show that dAT is 95% "open" while dTA is 90% "stacked" at 80 K (see Tab. 2).

Ab initio calculations on the stacking of free base pairs indicate that the stacking is primarily stabilized by dispersion attractions, which result from electron correlations [49,52]. The AMBER calculations do not consider electrons (although base stacking energies of free bases, followed as a function of twist and base separation, calculated by AMBER qualitatively match those predicted by *ab initio* calculations [53]). Therefore, higher-level theory may be needed to further understand how the 3′ base affects the conformational preference of these dinucleotides.

One final comment about the conformations of the dinucleotides concerns those with a guanine base. When G is in the 3′ position, it can form a "syn" type conformation (where G sits above the plane of the sugar) and hydrogen bond to the phosphate group. REMPI spectra on guanosines have shown that the "syn" conformation is present in the gas phase [54]. This provides an additional H-bond in the "open" conformers of dinucleotides with a 3′G and may be part of the reason why dTG favors the "open" conformation and why dAG "opens up" but dGA does not.

Table 2. Isomerization barriers (E_a, kcal/mol) determined from the ATD fits[a].

dinucleotide	K[b]	E_a[c]	
dAT	1:19	2.1	stacked → open[d]
dCT	1:80	—	
dGT	1:6	2.8	
dTT	1:50	—	
dTG	1:12	3.9	
dAA	3:1	2.6	open → stacked[d]
dCA	15:1	2.3	
dTA	9:1	3.2	
dAG	18:1	0.8	
dCG	5:2	4.1	
dGA	4:3	1.5	stacked → H-bond[d]
dGG	1:3	1.8	

[a]See text for explanation of ATD fits, [b]ratio of peaks in the 80 K ATDs (the shortest-time peak is listed first), [c]the uncertainty is ∼ 10%, [d]direction of isomerization that E_a corresponds to.

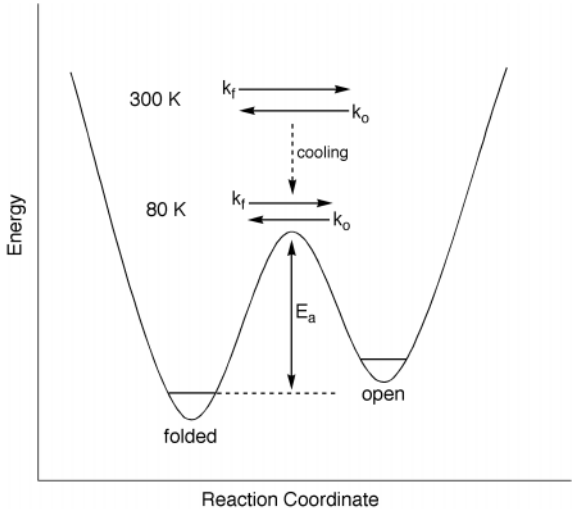

Fig. 6. Simple reaction coordinate diagram describing the isomerization of two different conformers. Labels are defined in the text.

3.2 Energetic properties of the dinucleotides

As shown in Figure 2, multiple peaks are observed in the low temperature ATDs but only one peak is present at higher temperatures. Since the multiple peaks at 80 K are known to be different conformations of the dinucleotide ions, the single peak at higher temperatures must be a result of the conformers rapidly isomerizing while they drift through the cell. If two or more different conformers rapidly interconvert while they travel through the drift cell, they will be detected at a single, averaged, arrival time (weighted by the relative amounts of each conformer) [35,43]. The ATD will thus show a single peak. A simple reaction coordinate diagram describing the interconversion of two conformers is shown in Figure 6.

At high enough temperatures, the average energy in the system will be greater than the isomerization barrier (E_a) between the two conformers. At this point, the two conformers can rapidly interconvert, yielding a single, time-averaged peak in the ATD. As the temperature is lowered, the average energy in the system decreases and approaches the barrier height. The isomerization process slows down and the ATD will show two peaks beginning to separate. When the energy drops below the isomerization barrier, the isomerization stops, the two conformers are essentially "frozen out", and the ATD shows two distinct peaks (the separation between the two peaks will depend on the cross-section difference of the two conformers). The relative intensities of the ATD peaks should reflect the actual conformer populations at this energy.

The height of the isomerization barrier, E_a, can be determined by measuring ATDs at different temperatures. The shape of the ATD peak(s) is dependent on how fast the two conformers isomerize, *i.e.*, their rate constants for interconversion (shown as k_f and k_0 in Fig. 6). The faster the conformers isomerize, the more the ATD peak resembles a single peak. As the isomerization slows down, the ATD begins to form two peaks. The shapes of the experimental ATDs are then fit with a theoretical model based on transport theory [55]. The only variables in this fit are the rate constants for the folded → open transition (k_f) and the open → folded transition (k_0). Examples of the fits are shown in Figure 7. Since the rate constants from the fits are known as a function of temperature, the barrier height can be determined from a simple Arrhenius analysis. Arrhenius plots of $\ln k$ *vs.* $1/T$ are shown for dTA^- and dAT^- in Figure 8. The barrier heights are obtained from the slopes of the lines and are collected for all of the dinucleotides in Table 2. The values vary between 1 and 4 kcal/mol and are not only base dependent but also sequence dependent.

The isomerization analyses for the three-conformer systems (dAC, dCC, dGC, and dTC) are more complicated. The problem is identifying how the conformers are connected. Does the "stacked" form directly convert into the "open" form or does it convert into the "H-bonded" form, which then converts into the "open" form? Do the "stacked" and "H-bonded" forms both convert into the "open" form? The analyses for all four 3′C dinucleotides have not been done at this time, but the experimental results for dCC indicate that this particular system may be divided into two separate 2-confomer systems.

The ATDs for dCC at different temperatures are shown in Figure 9. The "stacked" conformer separates first, between 400–500 K, while the "H-bonded" and "open" conformers separate between 100–200 K. Note that the time difference between the two peaks in the 300 K ATD, 18 μs, is in-between the time differences observed for the "stacked" and "H-bonded" conformers (12 μs) and for the "stacked" and "open" conformers (25 μs) in the ATDs given in Figure 3. This occurs because the peak at longest times in Figure 9 contains both the "H-bonded" and "open" conformers. Consequently, if significant amounts of the "H-bonded" conformers were present

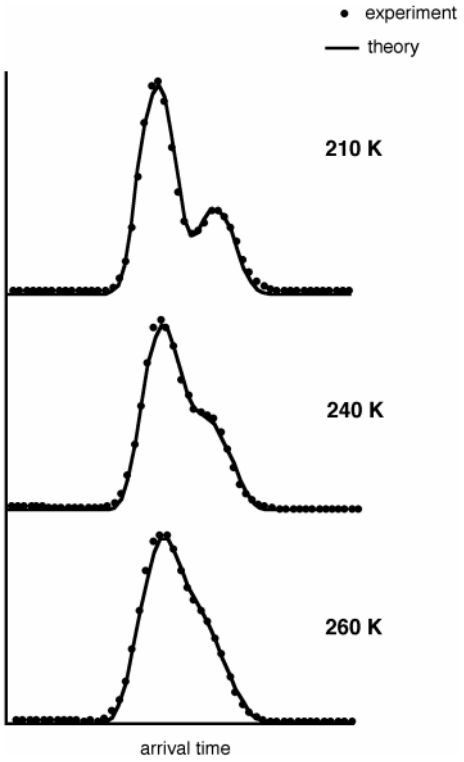

Fig. 7. ATD fits for dCG using the fitting procedure described in the text. Experimental data are shown as dots and the theoretical fits are shown as lines. The only variables used in the fits are the open → stacked and stacked → open rate constants.

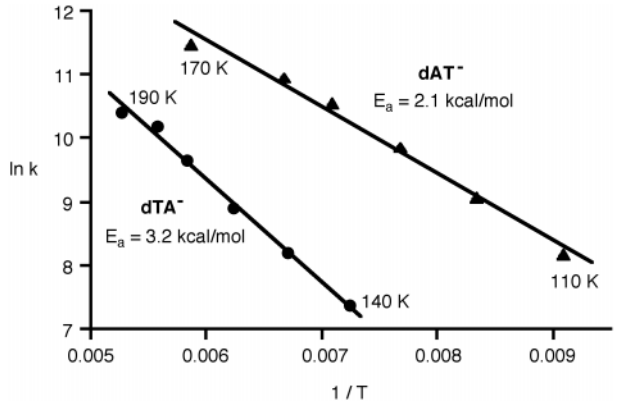

Fig. 8. Arrhenius plot for dTA (•) and dAT (▲) as determined from the ATD fits. The slope of the lines is proportional to the barrier height between the different conformations (see Fig. 6).

for the systems characterized by the ATD peaks shown in Figures 3a and 3b, those peaks would have been closer together. Hence, "H-bonded" conformers are not present for the dinucleotides in groups 1 and 2 (see Tab. 1).

The shorter-time peak in the 300 K dCC ATD is definitely the "stacked" form. The cross-section from this peak (132 ± 1 Å^2) agrees very well with the cross-section theoretically predicted for the "stacked" form (130 ± 2 Å^2). The cross-section obtained from the longer-time peak (146 ± 1 Å^2) is in-between the predicted "H-bonded" (142 ± 2 Å^2) and "open" (152 ± 2 Å^2) values. The ratio of the two peaks in the 300 K ATD is 1:3. The ratio of the three peaks in the 80 K ATD is 1:2:1. All of these results strongly indicate that the "stacked" form of dCC does not directly convert into the "open" form. Instead, the "stacked" form appears to convert into the "H-bonded" form, which then converts into the "open" form. This is schematically shown in the reaction coordinate diagram in Figure 10. Since the "H-bonded" ↔ "open" conversion occurs at lower temperatures than the "stacked" ↔ "H-bonded" conversion, the barrier between the "H-bonded" and "open" forms must be smaller than the barrier between the "stacked" and "H-bonded" forms.

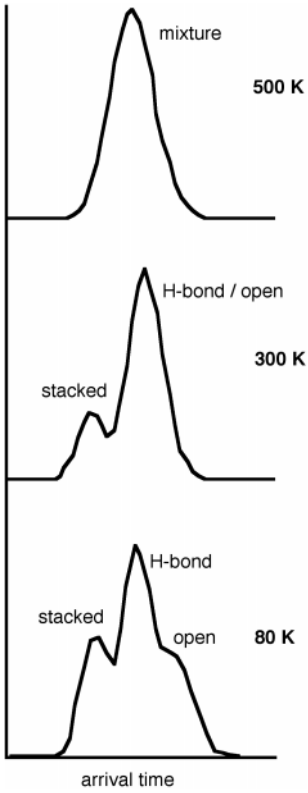

Fig. 9. ATDs for dCC at different temperatures. The "stacked" conformer separates from the others between 400–500 K. The "H-bonded" and "open" conformers separate between 100–200 K. The ratio of peaks in the 300 K ATD is 1:3, but at 80 K the peak ratios are 1:2:1, suggesting that the "stacked" conformer does not directly interconvert with the "open" conformer.

Because of the large temperature difference between the two transitions, each one can be effectively treated as an isolated system. In other words, it is assumed that the "stacked" conformer does not significantly affect the modeling of the H-bonded ↔ open transitions and the "open" conformer does not significantly affect the modeling of the stacked ↔ H-bonded transitions. Making these assumptions, the shapes of the temperature-dependent ATDs for dCC were fit with the same model as those used for the two-conformer systems described earlier. The Arrhenius plots obtained from the modeling are shown in Figure 11. For the open → H-bonded transition, the modeling yields

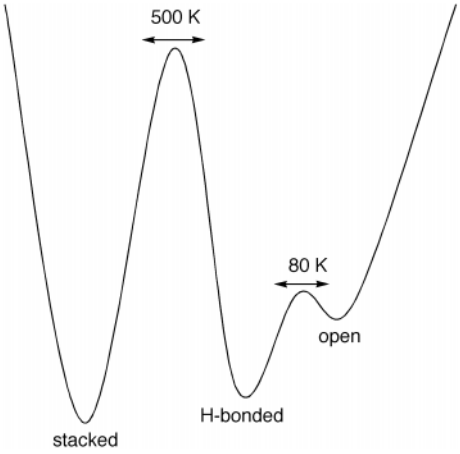

Fig. 10. Possible reaction coordinate diagram describing the isomerization of the "stacked", "H-bonded", and "open" forms of dCC. Based on the ATD data, the "H-bonded" and "open" conformers interconvert at low temperatures and thus have a small isomerization barrier. The "stacked" and "H-bonded" conformers isomerize at much higher temperatures and must have a larger barrier.

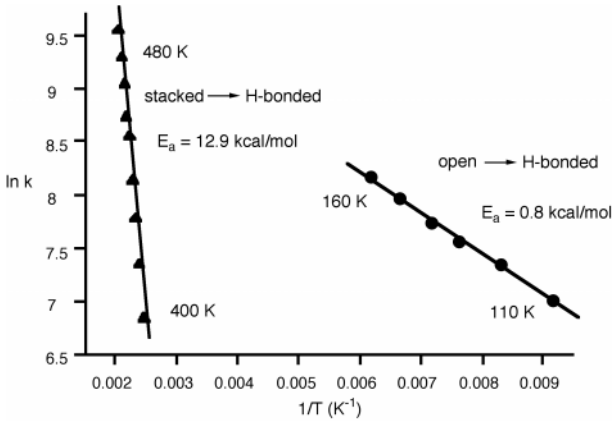

Fig. 11. Arrhenius plot for the interconversion of the "stacked" and "H-bonded" conformers (▲) and the "H-bonded" and "open" conformers (●) of dCC.

a barrier of 0.8 kcal/mol. For the stacked → H-bonded transition, the modeling yields a much higher barrier of 12.9 kcal/mol.

3.3 Temperature dependence of the dinucleotides

One final comment about the dinucleotides concerns their behavior at high temperatures. Figure 12 shows a plot of experimental cross-section, obtained from the ATDs, as a function of temperature for dAT and dTA. (The increase in cross-section at low temperature is due to the nature of the ion–He interaction potential as discussed in Section 3.1 and has been investigated for several systems [36,42]). Both of these dinucleotides have two peaks in their 80 K ATDs that correspond to the "stacked" and "open" conformers. However, the "open" form is favored (95%) for dAT and the "stacked" form is favored (90%)

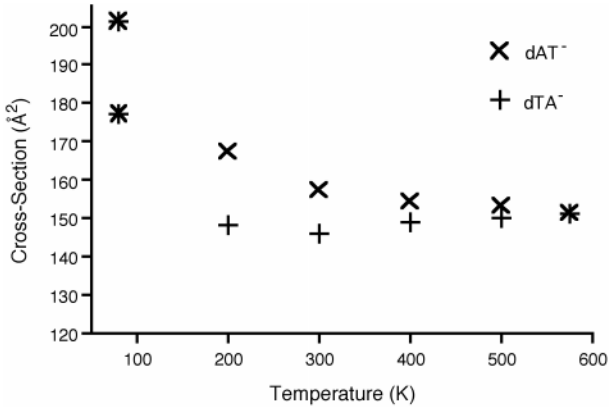

Fig. 12. Plot of experimental cross-section *vs.* temperature for dAT (×) and dTA (+). At 80 K, the "stacked" and "open" conformers are observed in the ATDs and so two cross-sections are represented on the plot. Note that the cross-sections for the "stacked" and "open" forms of dAT and dTA are the same at 80 K. Between 200–400 K dAT has a larger experimental cross-section than dTA, but at 580 K dAT and dTA have identical cross-sections.

for dTA. Table 2 lists the observed ATD peak ratios of the two conformers at 80 K.

In the plot in Figure 12, the two points at 80 K for each dinucleotide, due to the "stacked" and "open" conformers being separated in the ATD, are identical for dTA and dAT. That is, the "stacked" forms of dTA and dAT have the same cross-section and the "open" forms of dTA and dAT have the same cross-section at 80 K. By 200 K, the two conformers have begun to isomerize and the two peaks in the 80 K ATD have merged into a single peak in the 200 K ATD. However, between 200–400 K the experimental cross-section for dAT is significantly larger than that of dTA. Even though the "stacked" and "open" conformers are isomerizing while they drift through the cell, the average arrival time, and hence cross-section, is still weighted by the relative abundance of each conformer. According to the 80 K ATDs, dAT is 95% "open" (larger cross-section) while dTA is 90% "stacked" (smaller cross-section). Between 400–600 K, the cross-sections of dTA and dAT again merge to the same point. dAT has a slight decrease in cross-section in this temperature range (from 154 to 151 Å^2) but the cross-section of dTA reaches a minimum at 300 K (146 Å^2) and then increases to 151 Å^2 at 580 K, indicating dTA is "opening up" at the higher temperatures [36,44,56].

This trend can be seen more clearly by comparing the experimental cross-sections to the theoretically predicted values for the "stacked" and "open" conformers. The resulting cross-section *versus* temperature plots are shown in Figure 13. In these plots, the average cross-section of 10 "stacked" conformers and 10 "open" conformers (obtained from scatter plots analogous to Fig. 4) were calculated at different temperatures from 200 K to 600 K for comparison to experiment. For dAT, the experimental cross-section agrees fairly well with the theoretical value of the "open" form over the entire temperature range.

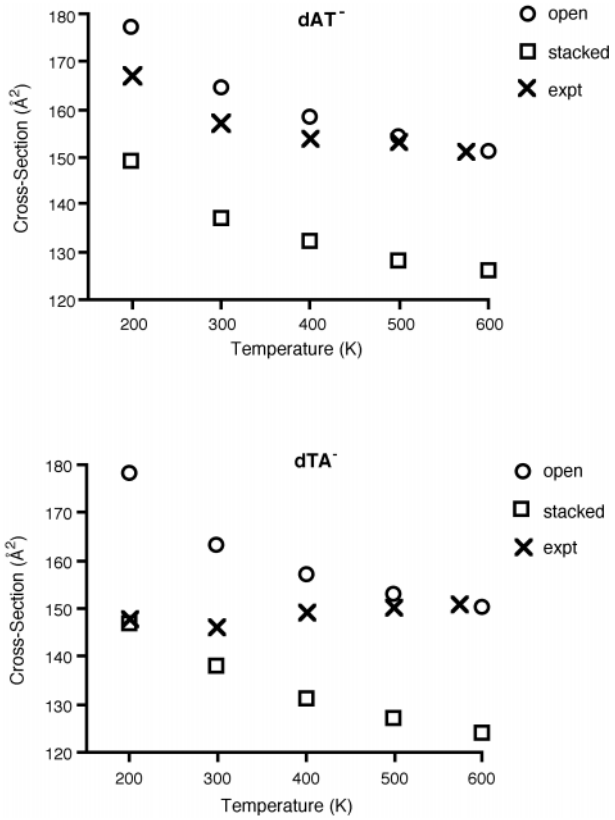

Fig. 13. Plots of cross-section *vs.* temperature for dAT and dTA. The crosses are experimental data, open circles are theoretical values for the "open" conformer, and the squares are theoretical values for the "stacked" conformer. The 80 K ATDs indicate that dAT is $\sim$ 95% "open" while dTA is $\sim$ 90% "stacked".

For dTA however, the experimental cross-section matches the "stacked" value at 200 K but as the temperature increases, the experimental cross-sections begin to agree with the "open" values. Similar trends were observed for all 16 dinucleotides. The experimental cross-sections for dAT, dCT, dGT, dTT, and dTG, which all favor the "open" form, slightly decreased from 400 K to 600 K, while the cross-sections of the other dinucleotides slightly increased over this temperature range (even dGA and dGG which did not show the "open" form in the 80 K ATDs).

A final point about these results should be made. In all systems, multiple conformations are observed at 80 K and reversible isomerization inevitably follows as the temperature in the drift cell is raised to 600 K. There is no evidence that "isolated" conformers are formed during the MALDI process. In this process the desorbed ions are formed moderately "hot" and subsequently are cooled in the drift cell. The qualitative mechanism is given in Figure 6. Hence, the hot ions sample all of the conformational space and the distribution observed in the ATDs is the result of the cooling process. The cross-sections of the various conformers observed in the ATDs all correlate quantitatively with low-energy conformers generated using our MM/MD calculations. There apparently are no "missing" conformers nor are the propensities of the observed conformers biased by the formation process.

4 Conclusions

Gas-phase conformational and energetic properties of 16 deprotonated dinucleotides were measured using ion mobility experiments and molecular modeling calculations. Three distinct families of conformers with similar energies but different collision cross-sections were identified– stacked, H-bonded, and open. At 80 K, these conformers can be individually observed in the experiments, but as the temperature is increased, they begin to interconvert as they drift through the mobility cell. The relative amounts of each conformer observed at 80 K and the temperature at which they begin to isomerize are base and sequence dependent. However, the 3′ base appears to play a major role in directing the conformational preferences of the dinucleotides. Theoretical modeling of arrival time distributions as a function of temperature allowed for the measurement of isomerization barrier heights, yielding values ranging from 0.8 to 12.9 kcal/mol (with most between 1–4 kcal/mol). At very high temperatures ($>$ 500 K), the dinucleotides appear to "open up" such that the bases are no longer stacked or hydrogen bonded to each other.

The support of the National Science Foundation under grants CHE9729146 and CHE0140215 is gratefully acknowledged.

References

1. R.R. Sinden, *DNA Structure and Function* (Academic Press, San Diego, 1994)
2. W. Saenger, *Principles of Nucleic Acid Structure* (Springer-Verlag, New York, 1984)
3. T.M. Alam, G.P. Drobny, Chem. Rev. **91**, 1545 (1991)
4. S.S. Wijmenga, B.N.M. van Buren, Prog. Nuc. Mag. Reson. Spectros. **32**, 287 (1998)
5. L. Zidek, R. Stefl, V. Sklena, Curr. Opin. Struct. Biol. **11**, 275 (2001)
6. O. Kennard, S.A. Salisbury, J. Biol. Chem. **268**, 10701 (1993)
7. F. Hillenkamp, M. Karas, R.C. Beavis, B.T. Chait, Anal. Chem. **63**, 1193A (1991)
8. J.B. Fenn, N. Mann, C.K. Weng, S.F. Wong, Mass Spectrom. Rev. **9**, 37 (1990)
9. P.A. Limbach, Mass Spectrom. Rev. **15**, 297 (1996)
10. E. Nordhoff, F. Kirpekar, P. Roepstorff, Mass Spectrom. Rev. **16**, 69 (1997)
11. B. Guo, Anal. Chem. **71**, 333R (1999)
12. S.A. McLuckey, S. Habibigourdarzi, J. Am. Chem. Soc. **115**, 12085 (1993)
13. D.P. Little, D.J. Aaaserud, G.A. Valaskovic, F.W. McLafferty, J. Am. Chem. Soc. **118**, 9352 (1996)
14. J.S. Ni, S.C. Pomerantz, J. Rozenski, Y. Zhang, J.A. McCloskey, Anal. Chem. **68**, 1989 (1996)
15. J.A. Loo, C.G. Edmonds, R.D. Smith, Science **248**, 201 (1990)

16. G. Tsaparilis, H. Nair, A. Somogyi, V.H. Wysocki, W. Zhong, J.H. Futrell, S.G. Summerfield, S.J. Gaskell, J. Am. Chem. Soc. **121**, 5142 (1999)
17. C.S. Hoaglund, Y. Liu, A.D. Ellington, M. Pagel, D.E. Clemmer, J. Am. Chem. Soc. **119**, 9051 (1997)
18. M.T. Bowers, P.R. Kemper, G. von Helden, P.A.M. van Koppen, Science **260**, 1446 (1993)
19. D.E. Clemmer, M.F. Jarrold, J. Mass Spectrom. **32**, 577 (1997)
20. G. von Helden, M.T. Hsu, N.G. Gotts, M.T. Bowers, J. Am. Chem. Soc. **97**, 8182 (1993)
21. N.G. Gotts, G. von Helden, M.T. Bowers, Int. J. Mass Spectrom. Ion Process. **149/150**, 217 (1995)
22. K.B. Shelimov, D.E. Clemmer, R.R. Hudgins, M.F. Jarrold, J. Am. Chem. Soc. **119**, 2240 (1997)
23. A.E. Counterman, D.E. Clemmer, J. Am. Chem. Soc. **123**, 1490 (2001)
24. T. Wyttenbach, G. von Helden, M.T. Bowers, J. Am. Chem. Soc. **118**, 8355 (1996)
25. T. Wyttenbach, J.J. Batka, J. Gidden, M.T. Bowers, Int. J. Mass Spectrom. **193**, 143 (1999)
26. T. Wyttenbach, J.E. Bushnell, M.T. Bowers, J. Am. Chem. Soc. **120**, 5098 (120)
27. B.S. Kinnear, D.T. Kaleta, M. Khotani, R.R. Hudgins, M.F. Jarrold, J. Am. Chem. Soc. **122**, 9243 (2000)
28. B.S. Kinnear, M.R. Hartings, M.F. Jarrold, J. Am. Chem. Soc. **123**, 5660 (2001)
29. A.E. Counterman, D.E. Clemmer, J. Phys. Chem. B **105**, 8092 (2001)
30. D.E. Clemmer, R.R. Hudgins, M.F. Jarrold, J. Am. Chem. Soc. **117**, 10141 (1995)
31. E.R. Bateman, C.S. Hoaglund-Hyzer, D.E. Clemmer, Anal. Chem. **73**, 6000 (2001)
32. S. Lee, T. Wyttenbach, M.T. Bowers, Int. J. Mass Spectrom. Ion Process. **167/168**, 605 (1997)
33. T. Wyttenbach, M. Witt, M.T. Bowers, Int. J. Mass Spectrom. **183**, 243 (1999)
34. T. Wyttenbach, M. Witt, M.T. Bowers, J. Am. Chem. Soc. **122**, 3458 (2000)
35. J. Gidden, J.E. Bushnell, M.T. Bowers, J. Am. Chem. Soc. **123**, 5610 (2001)
36. G. von Helden, T. Wyttenbach, M.T. Bowers, Int. J. Mass Spectrom. Ion Process. **146/147**, 349 (1995)
37. P.R. Kemper, M.T. Bowers, J. Am. Soc. Mass Spectrom. **1**, 197 (1990)
38. R.P. Morgan, J.H. Beynon, R.H. Bateman, B.N. Green, Int. J. Mass Spectrom. Ion Process. **28**, 171 (1978)
39. E.A. Mason, E.W. McDaniel, *Transport Properties of Ions in Gases* (Wiley, New York, 1988)
40. J. Gidden, T. Wyttenbach, J.J. Batka, P. Weis, A.T. Jackson, J.H. Scrivens, M.T. Bowers, J. Am. Soc. Mass Spectrom. **10**, 883 (1999)
41. D.A. Case, D.A. Pearlman, J.W. Caldwell, T.E. Cheatham III, W.S. Ross, C.L. Simmerling, T.A. Dearden, K.M. Merz, R.V. Stanton, A.L. Cheng, J.J. Vincent, M. Crowley, V. Tsui, R.J. Radner, Y. Duan, J. Pitera, I. Massova, G.L. Seibel, U.C. Singh, P.K. Weiner, P.A. Kollman, AMBER 6.0, University of California, San Francisco, 1999
42. T. Wyttenbach, G. von Helden, J.J. Batka, D. Carlat, M.T. Bowers, J. Am. Soc. Mass Spectrom. **8**, 275 (1997)
43. J. Gidden, A.T. Jackson, J.H. Scrivens, M.T. Bowers, Int. J. Mass Spectrom. **188**, 121 (1999)
44. J. Gidden, A.T. Jackson, J.H. Scrivens, M.T. Bowers, J. Am. Chem. Soc. **122**, 4692 (2000)
45. J. Sponer, J. Leszczynski, P. Hobza, Biopolymers **61**, 3 (2002)
46. D.M. Cheng, R.H. Sarma, J. Am. Chem. Soc. **100**, 7333 (1977)
47. P. Cieplak, P.A. Kollman, J. Am. Chem. Soc. **110**, 3734 (1987)
48. J. Norberg, L. Nilsson, Biophys. J. **74**, 394 (1998)
49. P. Hobza, J. Sponer, Chem. Rev. **99**, 3247 (1999)
50. E. Nir, K. Kleinermanns, M.S. deVries, Nature **408**, 949 (2000)
51. J. Norberg, L. Nilsson, J. Am. Chem. Soc. **117**, 10832 (1995)
52. J. Sponer, J. Leszczynski, P. Hobza, J. Phys. Chem. **100**, 5590 (1996)
53. P. Hobza, M. Kabelac, J. Sponer, P. Mejzlik, J. Vondrasek, J. Comput. Chem. **118**, 1136 (1997)
54. E. Nir, P. Imhoff, K. Kleinermanns, M.S. de Vries, J. Am. Chem. Soc. **122**, 8091 (2000)
55. I.R. Gatland, Case St. At. Phys. **4**, 369 (1974)
56. T. Wyttenbach, G. von Helden, M.T. Bowers, Int. J. Mass Spectrom. Ion Process. **165/166**, 377 (1997)