Comparison Between Solution Phase Stability and Gas Phase *Kinetic* Stability of Oligodeoxynucleotide Duplexes

Valérie Gabelica and Edwin De Pauw

Laboratoire de Spectrométrie de Masse, Bât. B6c, Université de Liège, B-4000 Liege, Belgium

Corresponding author : Valérie Gabelica, Lab. Mass Spectrom., Chemistry Institute B6c, ULG, B-4000 Liege, Belgium. Tel: +32 4 366 3432. Fax: +32 4 366 3413. E-mail: v.gabelica@ulg.ac.be

Running title: KINETIC STABLITY OF DUPLEX DNA IN VACUO

Number of pages : text : 18 ; figures : 6 ; tables : 3

ABSTRACT

The relative kinetic stabilities of different 16-mer oligonucleotide duplexes have been investigated by source-CID in a heated capillary electrospray ion source. They are compared to the relative stabilities in solution obtained by thermal denaturation monitored by UV-spectrophotometry. Results clearly show that both hydrogen bonding and base stacking interactions that are present in solution are maintained in the gas phase. This suggests that the electrospray process does preserve the double-helix structure of the DNA. A step by step opening of the double helix structure is proposed for the gas phase dissociation, competing with the covalent bond cleavage of bases. We also draw the attention to the fact that by source-CID, it is the *kinetic* stability of the complexes that is probed. In particular, this implies that only complexes of the same size can be compared.

Keywords: Electrospray, DNA, Non-covalent complexes, Source-CID, Denaturation

INTRODUCTION

The study of non-covalent complexes by electrospray ionization mass spectrometry (ESI-MS) is a constantly growing area of research. A large variety of synthetic[1,2] (cation-macrocycle, supramolecular assemblies,…) and biologically-relevant systems[3-5] (protein-protein, protein-ligand, protein-DNA, protein-RNA associations, etc…) have been studied to date. There are mainly two ways of using ESI-MS to study non-covalent complexes. First, full scan simple MS spectra with soft ionization conditions can be used to probe the composition of the solution (stoichiometry of the complex(es) and relative abundance(s) of the species). Second, one can also investigate the behavior of a specific complex or of a family of similar complexes in the gas phase by monitoring the dissociation when the internal energy is varied. The internal energy can be modulated mainly by thermal heating or by collisional activation in the electrospray source or in MS/MS experiments. Information can then be inferred about the complexes isolated in the absence of solvent. Whether the information obtained in the gas phase will be correlated to that in solution depends on the role of the solvent in the stabilization of the complex, and on the effect of the electrospray process on the conformation of the complex that is transferred from the solution to the gas phase.

Supramolecular complexes of biochemical interest often involve different kinds of interactions[6,7]: electrostatic (ion-ion, ion-dipole, dipole-dipole), hydrogen bonding, Van der Waals forces, hydrophobic interactions (strongly solvent-dependent). It is not straightforward to determine for a given complex if the gas phase behavior will resemble that in solution. It is therefore useful to investigate in detail some model systems involving only one or two types of interactions at a time.

We have chosen to study oligonucleotide duplexes as a model system. Three kinds of interactions are involved in the formation of the DNA double helix[8-10]. Phosphate groups are negatively charged, and electrostatic interactions, strongly dependent on the ionic strength of the solution, are important: the concentration of counter-ions needs to be high enough to counterbalance the charges of the phosphates, unless the two strands would separate because of charge repulsion. The purine and pyrimidine bases interact with each other by hydrogen bonding: this is the so-called Watson-Crick base pairing (fig. 1). GC-rich sequences are more stable than AT-rich ones. Finally, adjacent bases are also interacting with each other electrostatically (dipole-dipole short range interactions) and these interactions are favored in the double helix for conformational reasons. This kind of interaction is called 'base stacking'[11] and is sequence-dependent.

The first observations of DNA duplex structures by mass spectrometry date back to 1993[12,13] and the specificity of their formation has been discussed in 1995 by Ding and Anderegg[14]. In 1998, Williams and al. have investigated small model oligonucleotide duplexes using the BIRD method[15]. They correlated the stability of the duplexes in the gas phase with the number of hydrogen bonds between the strands, thereby inferring that the solution structure is in that respect maintained in the gas phase. This article reports the investigation of the gas phase kinetic stability of oligonucleotide duplexes by modulation of the internal energy in the electrospray source (source-CID), and the comparison of these results with the thermodynamic data in solution.

EXPERIMENTAL

Oligonucleotide single strands were chosen purposely to be non self-complementary and of significantly different masses to allow unambiguous interpretation of the spectra. Symmetrical base sequences were chosen to avoid the formation of homodimers, and to minimize mismatch base pairing. Oligonucleotides were purchased from Eurogentec (Sart-Tilman, Belgium), annealed in 0.1 M ammonium acetate to form duplex species which were desalted with MICROCON3 filters (Amicon, Beverly, MA) when the necessity was revealed by the mass spectra. The duplex solutions were 10 µM in a 80/20 (100 mM aqueous NH_4Ac/methanol) solution, and were injected by direct infusion at 2 µL/min.

Mass spectrometric measurements were performed using a Finnigan LCQ ion trap instrument (ThermoQuest, Bremen, Germany). The spray voltage was set to -3.8 kV (negative ion mode). The tube lens offset value has a great influence on the focalization and on the acceleration of the ions in the source, and thus on the internal energy the ions are given. This parameter was kept constant throughout the presented work: a value of –40 V was chosen, which was the best compromise for keeping the duplex intact while maintaining a good signal intensity. For each duplex, full scan mass spectra were recorded, varying the capillary-skimmer voltage (V_{CS} = 0 to –135 V) at fixed capillary temperature. The V_{CS} voltage is applied on the capillary exit, and the skimmer is at ground. Due to the inertia of the capillary temperature, it is easier to vary V_{CS} at fixed temperature than to vary T_{cap} at fixed voltage. It will be shown below that both are equivalent ways of giving internal energy to the ion. In order compare duplexes involving the same electrostatic repulsion, we always monitored the dissociation of $Duplex^{6-}$ species into their constitutive strands A^{3-} and B^{3-}. Like in reference [16], the percentage of duplex in each mass spectrum was calculated as follows :

$$\%\text{Duplex} = \frac{I_{\text{Dup}^{6-}}}{I_{\text{Dup}^{6-}} + \frac{\left(I_{A^{3-}} + I_{B^{3-}}\right)}{2}} \cdot 100\% \qquad (1)$$

Thermal denaturations in solution were performed using a Lambda 5 UV spectrophotometer (Perkin Elmer, Norwalk, CT, USA) interfaced with a PC XT for data acquisition. The quartz cells were heated by a Julabo F25 heating circulator (Kutztown, PA, USA). These measurements were performed with duplex concentrations of 2.5 µM, in order to have an absorbance range between 0.2 and 0.8. The concentrations were carefully set to the same value for the different UV-melting experiments. By doing so, the relative T_m values directly indicate the relative stability of the analyzed species (otherwise the T_m's would also depend on the concentration[17]).

RESULTS AND DISCUSSION

Hydrogen bonding

In order to investigate the lone contribution of hydrogen bonding to the gas phase kinetic stability of the DNA double helix, we performed source-CID experiments on a series of 16-mer duplexes (same size) with varying percentages of GC base pairs (varying number of H-bonds, see Figure 1). Table 1 summarizes the sequences of the studied duplexes, and the main results. As base stacking effects on stability are smaller than hydrogen bonding ones, their influence will be hidden in the series of comparisons of this first section. Moreover, all the duplexes have the same purine/pyrimidine sequence, and only 4 bases have been changed at a time. Figure 2 shows the solution thermal denaturation experiments on duplexes **1-5**, containing different percentages of GC base pairs. In solution, a temperature increase induces the denaturation of the duplex, which manifests itself in a hyperchromism measured by UV-spectrophotometry (single strands have a higher extinction coefficient ε than the duplex). The melting temperature T_m is defined as the temperature at which half the duplex is denatured, or melted. The higher the melting temperature, the greater the stability of the duplex in the solution.

Figure 3 shows the evolution of the duplex percentage in the spectra as a function of the capillary voltage V_{CS}, for duplexes **1-4** (containing 0% to 75% of GC base pairs), at two different capillary temperatures. These four duplex species are dissociating into their constitutive single strands exclusively. No covalent fragmentation is observed. In the case of the duplex **5** with 100% GC (the more stable one), extensive breaking of covalent bonds occurs before breaking of all the hydrogen bonds. Resulting spectra are highly noisy, and no sequence-informative fragments can be sorted out. We therefore show no dissociation curve for that particular duplex. In the text, covalent bond breaking will be called "fragmentation" and non-covalent breaking of the duplex into its single strands will be called "dissociation". Figure 3 clearly shows that the gas phase kinetic stability of a duplex is dramatically dependent on the number of hydrogen bonds between the two strands. By analogy with the melting temperature T_m, we can define a melting voltage V_m, as the voltage at which half the duplex is fragmented. The V_m values obtained are also summarized in Table 1. Similar approaches have already been described for source-CID [18] (V_m was called Vc_{50}), or MS/MS [19,20] (the collision energy of half fragmentation was called E_{50} or $E_{1/2}$).

Base stacking

The other type of interaction that is partly responsible for the stability of the DNA double helix is the "base stacking". In order to study the effect of this type of interaction independently from the others, we compared the behavior of 16-mer duplexes, all with 50% GC, but with different base sequences (Table 2). Figure 4 presents the source-CID experiment performed at 160°C. That capillary temperature was chosen in order to have a melting voltage of about 65 V, just in the middle of the voltage range available. We see in Figure 4 and Table 2 that duplex **7** (V_m = 71 V) is slightly more kinetically stable than duplex **6** (V_m = 65 V). Due to the impreciseness on the measured temperature in the thermal denaturation experiments with our equipment, no comparison with the T_m measurements is shown. However, a correlation can be made between the V_m values and thermodynamic quantities calculated by empirical nearest-neighbor models[10]. Nearest-neighbor models are established to calculate the sequence-dependent melting stability of short duplex DNA oligomers (20 base pairs or less). They are based on the statistical analysis of denaturation data and provide tables that can be used to calculate thermodynamic parameters for any sequence by an incremental method. We used the tables given by Sugimoto et al.[21] to calculate the duplex melting enthalpy ($\Delta H_{n\text{-}n}$) for **6** and **7**. These calculated values are summarized in Table 2.

Influence of the size of the complex

To emphasize that care should be taken when trying to compare gas phase kinetic data with solution phase thermodynamics, we present here a striking example with duplex species. We compared the behavior of duplex **3** (50% GC) with that of the same duplex with thymine bases added at both ends of strand A to form duplex **8** (see sequences in Table 3). The thymine bases have no partners to form hydrogen bonds with. They just increase the size, the mass and the number of degrees of freedom of duplex **8**, but do not change the energy of the interaction between the strands. The T_m values are the same in solution (54°C). However, in the source-CID experiments (Figure 5), there is a great difference between the V_m values of these two duplexes: the larger duplex has the greater V_m value.

DISCUSSION

On the source-CID method

Based on arguments developed recently[22,16], we can assume that the acceleration of the ions in the electrospray source increases the internal temperature of the ion in a thermal way, and thus that increasing V_{CS} has exactly the same effect as increasing T_{cap}. Source-CID was thus described as a "thermal denaturation in the gas phase"[16]. This is also supported by this study. Figure 6 summarizes the V_m values that have been obtained for duplexes 1-4 at different capillary temperatures. If more internal energy is already given by thermal heating in the capillary, less is remaining to be added by acceleration in order to achieve 50% of fragmentation. The linear relationship between T and V indicates that the two activation methods are equivalent, and that acceleration in the medium pressure region produces a slow heating of the ions: increasing V_{CS} by 10 V has the same effect as increasing the capillary temperature by about 10°C. Changing the capillary temperature just shifts the internal energy range available when varying the V_{CS} from the minimum to the maximum value (0 to 135 V). This was applied in the study of base stacking effects: a capillary temperature of 160°C was chosen to observe the fragmentation in the middle of the V_{CS} range.

The source-CID method has some advantages for the study of non-covalent complexes. Its inability to mass select a parent ion is usually not a handicap because the weakly associated non-covalent species are usually the ones with the lowest fragmentation barriers. A rapid energy-dependent study can be performed by recording the full scan mass spectra as a function of the acceleration voltage. Figure 4 nicely shows that source-CID is a sensitive method to probe subtle energetic differences between complexes.

The use of source-CID and of MS/MS to study non-covalent complexes in the gas phase has already been reported before, and parameters like Vm^{16}, $Vc_{50}{}^{18}$, $E_{50}{}^{19}$, $E_{1/2}{}^{20}$ were derived to characterize, according to the authors, the *stability* of the complexes in the gas phase. The term *stability* is misused in this case, and the term *"kinetic stability"* (the contrary of *lability*) should be used instead, because in source-CID, as well as of traditional MS/MS, it is a reaction *rate* that is reflected by the fragmentation ratio[23]. Shortly, (1) the observed fragmentation ratio depends on the time window of the experiment and on the dissociation rate constant k, (2) the rate constant k is a function of the internal energy E of the ions and (3) the k(E) function depends on the dissociation barrier E_0, but also on the vibration frequencies of the reactant and of the transition state. In other words, k(E) depends on the reactant (its size and its whole set of bonds) and on the fragmentation pathway. Usually, the internal energy of the ions is not known and one can thus not derive absolute E_0 values (an exception is the case of BIRD experiments). One can only work by comparisons, but the problem is: "What can be compared to what?". Considering the points mentioned before, the relative V_m values will reflect the relative E_0 values only when comparing complexes of the same size, same structure, and fragmenting the same way. The influence of the size of the complex observed in Figure 5 can easily be explained considering the simple RRK formula for unimolecular dissociation (Equation 2)[23].

$$k(E) = \nu_s \cdot \left(\frac{E - E_0}{E} \right)^{DOF-1} \tag{2}$$

This is a crude approximation of k(E), but the equation makes the following qualitative reasoning easy to understand. The larger the number of degrees of freedom (DOF), the larger the probability for a given total energy E to be redistributed on other normal modes than the reaction coordinate, the slower the fragmentation reaction. In the case of duplex **8**, it required a higher acceleration voltage (a higher energy) for the same reaction rate as for duplex **3** to be achieved. When the dramatic influence of the size of the complex on reaction rate constants is forgotten, it can lead to somewhat hasty conclusions.

Duplex DNA in solution and in the gas phase

In this work, the kinetic stabilities of complexes of same size and structure, and fragmenting the same way (Duplex^{6-} □ $A^{3-} + B^{3-}$) are compared. This ensures that the variation of the fragmentation rate should only be due to a variation in the activation barrier E_0. The variation of activation energy as a function of the number of hydrogen bonds between DNA strands has already been studied by BIRD[15] and by MS/MS[20]. The phenomenological trend is confirmed by this source-CID study on duplexes **1-4**: the activation barrier E_0 depends on the percentage of GC base pairs in the DNA strands (Figure 3 and Table 1). This parallels the stability scale in solution, where T_m depends strongly on the number of hydrogen bonds between the strands. Comparing the stability in solution and the kinetic stability in the gas phase for duplexes **6-7** (Figure 4 and Table 2), we observed for the first time the effect of base stacking interactions in the gas phase. This study suggests that the double-helix structure of DNA that is present in solution is maintained during the electrospray process that gently transfers the ions from the solution to the gas phase, and even subtle energetic differences present in solution, and strongly conformation-dependent, can be reflected in the gas phase.

On the fragmentation mechanism

In the source-CID experiments, the simple dissociation of the duplex into its constitutive strands is the main fragmentation pathway observed, except for duplex **5**, for which extensive covalent bond cleavage occurs before the rupture of all the hydrogen bonds. The threshold for covalent bond cleavage of duplex **5** (see Figure 6) limits the range of E_0(dissociation) that are accessible for probing by source-CID, because another reaction pathway is competing with the dissociation. In the gas phase, the combination of multiple weak interactions can make a non-covalent complex more kinetically stable than covalent bonds under some activation conditions[24].

In solution, for short oligomers like those studied here, thermal denaturation experiments are usually analyzed considering a simple two-step mechanism (duplex ⇔ single strands)[25]. The T_m values are the temperatures at which half the duplex is separated in single strands. The two-step mechanism provides a useful framework for the interpretation of the data, in the gas phase as well as in solution. If the dissociation of the duplex involves the simultaneous breaking of all the hydrogen bonds, the activation barrier E_0 depends on the total number of H-bonds involved. The covalent bond cleavage of the duplex is another reaction channel that can compete when the E_0 for dissociation is too high. However, such all-or-none model is too simplified to be realistic. This would imply the simultaneous breaking of 32 to 44 hydrogen bonds, which is a highly improbable event involving a very high activation barrier. The separation of the strands would more likely imply a progressive unzipping of the strands, and would only require successive low barriers steps. This mechanism is compatible with a continuous supply of small quantities of energy by slow heating, which favors the dissociation of the duplex.

ACKNOWLEDGEMENTS

We would like to thank F. Rosu (Biospectroscopy Laboratory, ULg) for having performed the thermal denaturation experiments by UV spectrophotometry. V.G. is grateful to the F.N.R.S. (Belgium) for the PhD fellowship. This work has been supported by the "Actions de Recherche Concertée (ARC)" (Direction de la Recherche Scientifique – Communnauté Française de Belgique).

REFERENCES

1) Vincenti M. *J. Mass Spectrom.* 1995; **30**: 925.

2) Schalley CA. *Int. J. Mass Spectrom.* 2000; **194**: 11.

3) Smith DL, Zhang Z. *Mass Spectrom. Rev.* 1994; **13**: 411.

4) Loo JA. *Mass Spectrom. Rev.* 1997; **16**: 1.

5) Smith RD, Bruce JE, Wu Q, Lei QP. *Chem. Soc. Rev.* 1997; **26**: 191.

6) Lehn J-M. *Supramolecular chemistry,* VCH Verlagesellschaft (1995).

7) Schneider H.-J. *Angew. Chem. Int. Ed. Engl.* 1991; **30**: 1417.

8) Sharp KA, Honig B. *Curr. Opin. Struct. Biol.* 1995; **5**: 323.

9) Turner DH. *Curr. Opin. Struct. Biol.* 1996; **6**: 299 .

10) Owczarzy R, Vallone PM, Gallo FJ, Paner TM, Lane MJ, Benight AS. *Biopolymers* 1998; **44**: 217.

11) Sponer J, Laszczynski J, Hobza P. *J. Biomol. Struct. Dyn.* 1996; **14**: 117.

12) Ganem B, Li Y-T, Henion JD. *Tetrahedron Lett.* 1993; **34**: 1445 .

13) Light-Wahl KJ, Springer DL, Winger BE, Edmonds CG, Camp DG, Thrall BD, Smith RD. *J. Am. Chem. Soc.* 1993; **115**: 803.

14) Ding J, Anderegg RJ. *J. Am. Soc. Mass Spectrom.* 1995; **6**: 159 .

15) Schnier PD, Klassen JS, Strittmatter EF, Williams ER. *J. Am. Chem. Soc.* 1998; **120**: 9605.

16) Gabelica V, Rosu F, Houssier C, De Pauw E. *Rapid Commun. Mass Spectrom.* 2000; **14**: 464.

17) Breslauer KJ. *Methods Enzymol.* 1995; **259**: 221.

18) Rogniaux H, Van Dorsselaer A, Barth P, Biellmann JF, Brabanton J, van Zandt M, Chevrier B, Howard E, Mitschler A, Potier N, Urzhumtseva L, Moras D, Podjarny A. *J. Am. Soc. Mass Spectrom.* 1999; **10**: 635.

19) Gao J, Wu Q, Carbeck JD, Lei QP, Smith RD, Whitesides GM. *Biophys. J.* 1999; **76**: 3253.

20) Wan KX, Shibue T, Gross ML. *J. Am. Soc. Mass Spectrom.* 2000; **11**: 450.

21) Hobza P, Sponer J. *Chem. Rev.* 1999; **99**: 3247.

22) Drahos L, Heeren RMA, Collette C, De Pauw E, Vékey K. *J. Mass Spectrom.* 1999; **34**: 1373.

23) Vékey K. *J. Mass Spectrom.* 1996; **31**: 445.

24) Penn SG, He F, Lebrilla CB. *J. Phys. Chem. B* 1998; **102**: 9119.

25) Cantor CR, Shimmel PR. In *Biophysical Chemistry*. *Part III: The Behavior of Biological Macromolecules.* Freeman and Company. San Fransisco. 1980; 1215.

FIGURE LEGENDS

Figure 1

Watson-Crick base pairing in the B-helix form of DNA. The hydrogen bonds are indicated by dashed lines. Thymine (T) and adenine (A) are bound by 2 H-bonds (top), whether cytosine (C) and guanine (G) are bound by 3 H-bonds.

Figure 2

Thermal denaturation curves of duplexes **1-5** in solution measured by UV-spectrophotometry. The absorbance ranges have been scaled from 0 to 1 (arbitrary units) for all duplexes.

Figure 3

Source-CID experiments on duplexes **1-4**, at two capillary temperatures (160°C and 180°C). The duplex percentage was calculated by equation (1) for each spectrum recorded at a different capillary-skimmer voltage (or V_{CS}).

Figure 4

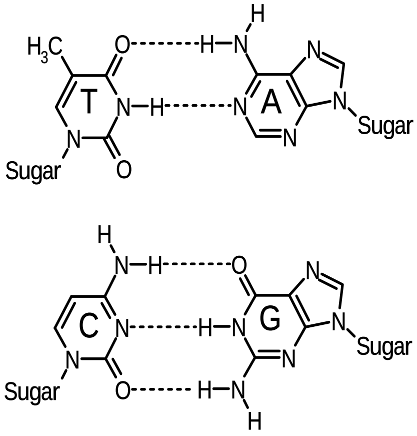
Source-CID experiments on duplexes **3** and **8** at $T_{cap} = 160°C$. The duplex percentages were calculated the same way as in figure 3.

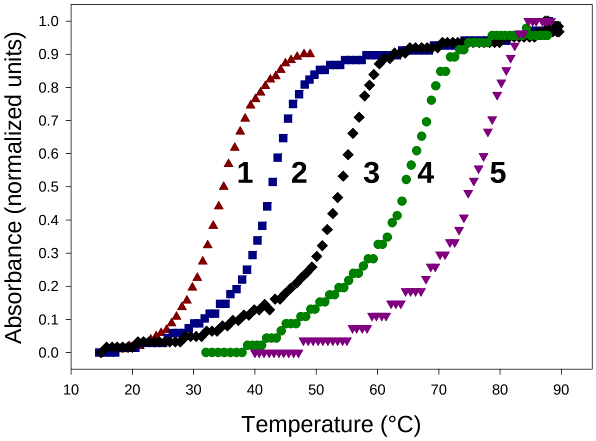
Figure 5

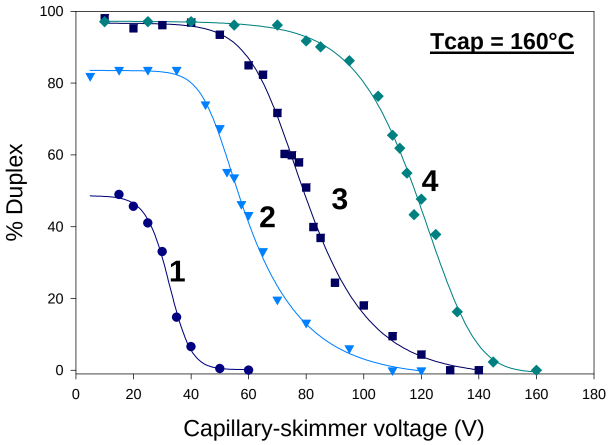
Source-CID experiments on duplexes **6-7** at $T_{cap} = 160°C$. The duplex percentages were calculated the same way as in figure 3.

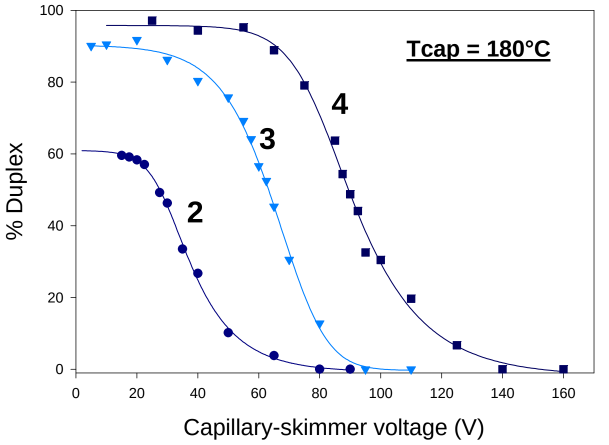
Figure 6

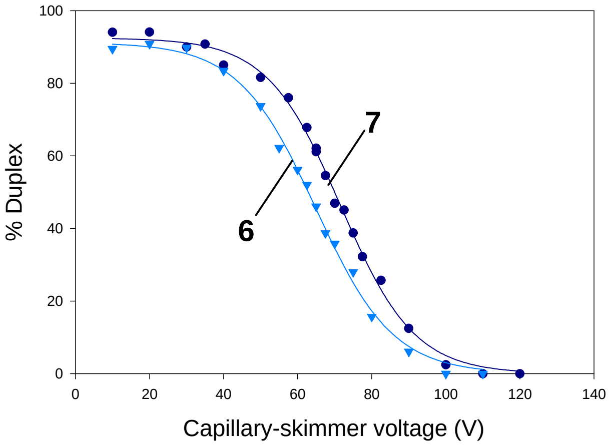
Voltages of half-fragmentation (V_m) for duplexes **1-4** at different capillary temperatures T_{cap}. For duplex **5**, the points represent the onset of the covalent bond fragmentation. These (V_m,T_{cap}) values can be considered as the upper limit for the source-CID method to be applicable.


T
A
C
G
Sugar
Sugar
Sugar
Sugar

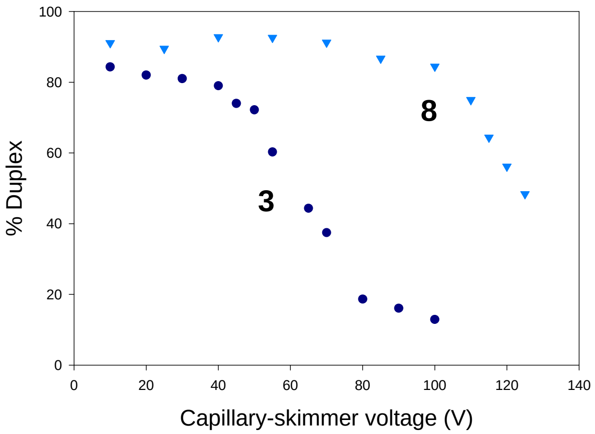

Absorbance (normalized units)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
10
20
30
40
50
60
70
80
90
Temperature (°C)
1
2
3
4
5


Tcap = 160°C
% Duplex
0
20
40
60
80
100
Capillary-skimmer voltage (V)
0
20
40
60
80
100
120
140
160
180
1
2
3
4


Tcap = 180°C
% Duplex
0
20
40
60
80
100
Capillary-skimmer voltage (V)
0
20
40
60
80
100
120
140
160
2
3
4



100
80
60
40
20
0
% Duplex
0
20
40
60
80
100
120
140
Capillary-skimmer voltage (V)
7
6


100
80
60
40
20
0
0
20
40
60
80
100
120
140
% Duplex
Capillary-skimmer voltage (V)
8
3

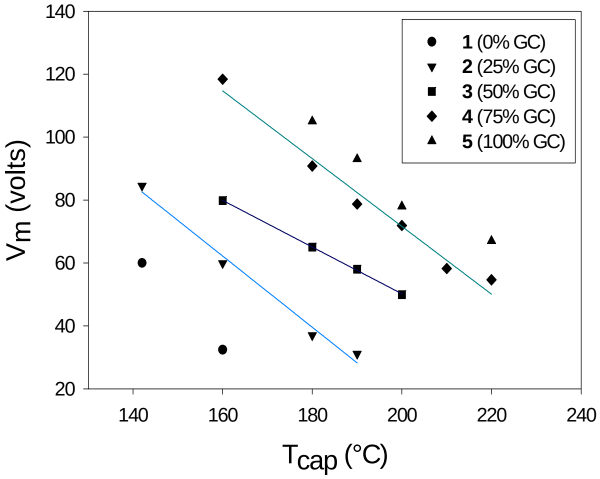

140
120
100
80
60
40
20
140
160
180
200
220
240
Vm (volts)
Tcap (°C)
● 1 (0% GC)
▼ 2 (25% GC)
■ 3 (50% GC)
◆ 4 (75% GC)
▲ 5 (100% GC)

Table 1

Summary of the study of the hydrogen bonding in duplex DNA : base sequences, GC percentage, measured values of T_m (temperature of half-denaturation in solution) and V_m (voltage of half-fragmentation in source-CID experiments) at two different capillary temperatures T_{cap} for duplexes **1-5**. Note: only the sequence of strand A (5'-3') is indicated. Strand B is the complementary of strand A.

	Base sequence	**% G≡C**	**T_m**	**V_m (T_{cap} = 160°C)**	**V_m (T_{cap} = 180°C)**
1	AAATTATAATATTAAA	0 %	34°C	32 V	-
2	GGATTATAATATTAGG	25 %	43°C	60 V	37 V
3	GGGCTATAATATCGGG	50 %	54°C	80 V	65 V
4	GGGCCGTAATGCCGGG	75 %	64.5°C	118 V	91 V
5	GGGCCGCGGCGCCGGG	100 %	76°C	-	> 105 V

Table 2

Summary of the study of the base stacking in duplex DNA : base sequences, GC percentage, and V_m (voltage of half-fragmentation in source-CID experiments) at T_{cap} = 160°C, and calculated duplex melting enthalpies ($\Delta H_{n\text{-}n}$) for duplexes **6** and **7**. Note: only the sequence of strand A (5'-3') is indicated. Strand B is the complementary of strand A.

	Base sequence	**% G≡C**	**V_m (T_{cap} = 160°C)**	**$\Delta H_{n\text{-}n}$ (kcal/mol)**
6	AGACTGTGAGTCAGTG	50 %	65 V	-123.8
7	GGGCTTTTAAAACGGG	50 %	71 V	-137.1

Table 3

Summary of the study of the influence of the complex size: complete base sequence (5'-3') and V_m values at T_{cap} = 160°C.

	Base sequence	**V_m (T_{cap} = 160°C)**
3	GGGCTATAATATCGGG CCCGATATTATAGCCC	68 V
8	TTTGGGCTATAATATCGGGTTT CCCGATATTATAGCCC	125 V