Specific and Nonspecific Dimer Formation in the Electrospray Ionization Mass Spectrometry of Oligonucleotides

Jianmei Ding and Robert J. Anderegg
Glaxo Research Laboratories, Research Triangle Park, NC 27709 North Carolina, USA

Specific and nonspecific noncovalent dimer ions of oligonucleotides (ODNs) were observed when mixtures of complementary or noncomplementary strands were analyzed via negative ion electrospray ionization mass spectrometry. Dimer formation was concentration dependent and nearly always occurred when the concentration of ODN exceeded 100 μM. Dimers were observed even for short-length ODNs for which the melting temperature \(T_m\) was well below the experimental temperature and which, therefore, would not be expected to form stable solution duplexes. The abundance of the heterodimer ions seems to correlate with the number of expected hydrogen bonds from Watson–Crick base pairing. As the energy of the incoming ion beam (orifice potential) was increased, the absolute and relative abundance of the dimer ions unexpectedly increased. ([Am Soc Mass Spectrom 1995, 6, 159–164])

The development of electrospray ionization mass spectrometry (ESI-MS) has opened up vast new applications for mass spectrometry in the study of biopolymers [1–3]. In addition to molecular mass information of very large molecules and primary sequence data, ESI-MS holds the potential to probe a variety of conformational questions and even to explore the interaction of proteins with other molecules. One of the most exciting recent observations has been that noncovalent complexes of proteins and small molecules or of proteins and other proteins could be observed in ESI mass spectra [4–16]. This observation led to the speculation, yet to be convincingly demonstrated, that the abundance of the ions that represent these noncovalent complexes should be related in a predictable way to the binding strength of the complex in solution. Since the original report [4], many noncovalent complexes have been described, including protein–ligand interactions [5–7, 10, 14, 15], protein dimers [9, 11, 14], ternary complexes of multiple proteins–ligands [8, 16], and oligonucleotides [12, 13].

One difficult aspect of these studies is the selection of appropriate control experiments to demonstrate unequivocally that the binding is specific. Nonspecific complexes are observed routinely in mass spectrometry via a variety of ionization methods, particularly as the concentration of analyte increases beyond what would be required to observe a comfortable signal–to–noise ratio. In our analysis of oligonucleotides, we have observed many noncovalent dimers, most of which could not be explained on the basis of Watson–Crick base pairing of complementary strands. We undertook this study to examine the effect of base sequence, concentration, and instrument variables on the production of ions that correspond to noncovalent complexes.

Experimental

All oligodeoxynucleotides (ODNs) were purchased from Research Genetics (Huntsville, AL) and were used without further purification. The absence of other ODN impurities was confirmed by mass spectrometry. The samples were dissolved in 10-mM aqueous ammonium bicarbonate at pH 8.5. The UV absorbance of each solution was measured at 260 nm via a Hewlett-Packard 8452A diode array spectrophotometer (Hewlett-Packard, Avondale, PA). The solution concentrations were calculated from the absorbance by using Beer’s law and were adjusted to 100 or 10 μM in ODN by the addition of buffer. The molar absorptivity for the Beer’s law calculation was computed from the program OLIQIO 4.04 (National Biosciences, Inc., Plymouth, MN). The solutions were not generally annealed before analysis.

The samples were infused into a Sciex (Thornton, Ontario) API-III mass spectrometer at 2 μL/min via a Harvard Model 22 syringe pump (Harvard Apparatus, South Natick, MA). The mass spectrometer was operated in the negative ion mode, with an orifice potential of –60 to –120 V. A mass range of 400–2390-u was scanned with a step size of 0.2 u and a dwell time of 2
Results and Discussion

The polyanionic backbone of oligonucleotides makes them readily observable by negative ion electrospray ionization mass spectrometry (ESI-MS). Figure 1 shows the ESI mass spectrum for the deoxyhexanucleotide \(d(GAGTTC)\) at a concentration of 100 \(\mu\)M in 10-mM aqueous ammonium bicarbonate at \(pH\) 8.5. The dominant ion is a doubly charged anion at \(m/z\) 902.5. Sodium replacement ions at \(m/z\) 11 intervals higher result from ubiquitous sodium contamination and corroborate the assignment of this cluster as doubly charged [18]. There is also a singly charged cluster beginning at \(m/z\) 1805.8, with sodium replacements at \(m/z\) 22 intervals, and a triply charged anion at \(m/z\) 601.1. Most interesting, however, is the appearance of a group of ions around \(m/z\) 1203.7, seen in the inset of Figure 1. The sodium replacement ions occur at a spacing of about \(m/z\) 7, which suggests that the ions are triply charged and implies a molecular mass of 3614, exactly double the predicted mass of the hexanucleotide. The presence of this species is confirmed by closer examination of the ions around \(m/z\) 1805. In addition to the \(m/z\) 22 sodium replacements mentioned previously, there is a series of ions spaced at \(m/z\) 11, which suggest that some of the ions in the cluster are doubly charged, not singly charged. The doubly charged species would indicate a molecular mass of 3614, in addition to the singly charged mass-1807 material.

The appearance of this apparent dimer of the hexanucleotide is concentration dependent. If the solution is diluted to 10 \(\mu\)M, the appearance of the spectrum is largely unchanged; the singly, doubly, and triply charged species remain, but the dimer ions disappear. If there were an impurity in the sample at twice the mass of the hexamer, it should persist regardless of concentration, so we conclude that the dimer is noncovalent. Such noncovalent dimers are observed in every ODN we have investigated if the concentration of nucleotide is greater than or equal to 100 \(\mu\)M. The sequence \(d(GAGTTC)\) is not self-complementary; only four of the six bases can form Watson-Crick base pairs between two strands of \(d(GAGTTC)\).

When mixtures of different ODN strands are prepared, the dimers are again observed, but can now form between like strands (homodimers) or between unlike strands (heterodimers). Figure 2a shows the mass spectrum of a mixture of \(d(GAGTTC)\) and its complementary sequence \(d(GAACTC)\), each at 100 \(\mu\)M. Figure 2b shows the mass spectra of various binary mixtures at a concentration of 100 \(\mu\)M per strand of ODN in 10-mM aqueous ammonium bicarbonate at \(pH\) 8.5. Sodium adduct ions are observed and were used to confirm charge state assignments. The inset is a magnification of regions a and b.
As before, the doubly charged monomers are the most abundant ions, with anions from both monomers in evidence, along with the triply charged counterparts. However, in the region of the spectrum where one would expect to find signals of triply charged dimers (Figure 2a, inset), ions appear that correspond to homodimers of \( d(GAGTTC) \) at \( m/z \) 1203, homodimers of \( d(GAACTC) \) at \( m/z \) 1183, and heterodimers of both at \( m/z \) 1193. As with the individual strands, when the concentration of the mixture is reduced by tenfold, ions from all the dimers disappear.

The relative abundances of the homo- and heterodimers can be predicted on the basis of the binomial expansion if there is no specific interaction between any two strands (i.e., the interaction is purely statistical). If the monomers were exactly equal in concentration and had exactly the same ionization efficiencies, one would predict a ratio of 0.25:0.5:0.25 for the homodimer:heterodimer:homodimer. Because the concentrations or ionization efficiencies were not precisely equal, we used the relative abundance of the ions that correspond to the monomers as a measure of relative concentration and predicted a ratio of 23:50:27 for this mixture. The observed ratio is 16:62:22, in reasonable agreement with the predicted ratio, which suggests that there is little specificity in the interaction.

In mixed sequence OONs, the ionization efficiencies between different monomers are roughly comparable. However, in sequences where one base predominates, for example \( d(T)_{10} \) versus \( d(G)_{10}\), there is a large difference in ion current obtained from solutions of equal concentration. The relative order of response is \( \text{poly-T} > \text{poly-C} \gg \text{poly-G} \). The response of \( d(T)_{10}\) is over twentyfold greater than the response for an equal concentration solution of \( d(G)_{10}\). This same discrimination is observed in the matrix-assisted laser desorption mass spectrometry of ODNs [20].

As a further control, we prepared a mixture of \( d(GAGTTC) \) and \( d(CTCAAG) \). The latter hexanucleotide has the same base composition as the complementary strand for \( d(GAGTTC) \), but with less possibility of Watson–Crick base pairing. The mass spectrum of that mixture is shown for comparison in Figure 2b. Again the monomers, homodimers, and heterodimers are observed in the statistically predicted ratio (see Table 1), which indicates nonspecific complex formation.

Normally, even complementary ODNs of short length have only limited ability to form stable duplexes in solution at room temperature. The stability of a duplex in solution is usually reported by its melting temperature (\( T_m \)), that is, the temperature at which half the duplex is dissociated [21]. The melting temperature can be predicted from the base sequence and is related to the extent of Watson–Crick base pairing possible. For the hexamer \( d(GAGTTC) \) with its complementary strand, both at a concentration of 100 \( \mu \text{M} \), \( T_m = -20.2 \, ^\circ \text{C} \) is predicted. At the temperature of our

### Table 1. Comparison of the observed relative abundance of homodimer and heterodimer ions to that predicted based on statistical combination

<table>
<thead>
<tr>
<th>Oligonucleotides</th>
<th>Number of H bonds</th>
<th>Predicted</th>
<th>Observed</th>
<th>Excess* (meas.-pred.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( d(GAGTTC) )</td>
<td>15</td>
<td>23</td>
<td>16</td>
<td>12</td>
</tr>
<tr>
<td>( d(GAACTC) )</td>
<td>50*</td>
<td>62</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>( d(GAGTTC) )</td>
<td>8</td>
<td>28</td>
<td>29</td>
<td>2</td>
</tr>
<tr>
<td>( d(GAACTC) )</td>
<td>50</td>
<td>48</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>( d(GAGTTC) )</td>
<td>18</td>
<td>23</td>
<td>13</td>
<td>19</td>
</tr>
<tr>
<td>( d(GAACTC) )</td>
<td>50</td>
<td>69</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>( d(GAGTTC) )</td>
<td>12</td>
<td>28</td>
<td>24</td>
<td>10</td>
</tr>
<tr>
<td>( d(GAGGAAA) )</td>
<td>50</td>
<td>60</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>( d(GAGTTC) )</td>
<td>35</td>
<td>24</td>
<td>7</td>
<td>17</td>
</tr>
<tr>
<td>( d(GAGGAAA) )</td>
<td>50</td>
<td>87</td>
<td>37</td>
<td>37</td>
</tr>
<tr>
<td>( d(GAGTTC) )</td>
<td>27</td>
<td>28</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>( d(GAGGAAA) )</td>
<td>50</td>
<td>80</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>

*The excess relative abundance between measured and predicted for the heterodimer.

*The number of H bonds from the best alignment of the two ODN strands.

*The bold font represents a pair of complementary strands.

*The predicted relative abundance ratio of the smaller homodimer \( 2M_1 \).

*The predicted relative abundance ratio of the heterodimer \( [M_1 + M_2] \).

*The predicted relative abundance ratio of the larger homodimer \( 2M_2 \).

*The normal font represents a pair of noncomplementary strands.
strands exceeds that predicted by a larger margin. One mixture of noncomplementary strands Mi = d(AAAAACA)

sents the hctc rod ime r.

d(TCTCTTTT) and served relative abundances for the hexamer and oc­
solutions are about 5% relative abundance. The ions

erodimers; see insets, Figure 3) in both the mixtures of

complementary and noncomplementary ODNs . The degree of specificity in the interaction of the het­
both the complementary and noncomplementary mix­

tadesamers, in both cases by using a pair of sequences

must be used. We explored ODN octamers and pen­

terodimer over the noncomplementary heterodimer, al­

To actually observe specific duplexes, longer ODNs

experiments, no stable duplex would be expected in

the solution from which the mass spectrum was ob­
tained. As a further check, the absorbance at 260 nm

was measured at 10 and 90 °C. No appreciable increase

in absorbance was noted, which indicates no duplex

was present. The Tm's are concentration dependent [21]

and increase for these hexanucleotides about 10 °C per

log increase in concentration. Even if the concentra­
tion increased a hundredfold in the desolvation of the

droplet during electrospray, the calculated Tm would

be below room temperature.

To actually observe specific duplexes, longer ODNs

must be used. We explored ODN octamers and pen­
tadesamers, in both cases by using a pair of sequences

that was complementary and a pair of the same base

composition that was not complementary as a control.

The ESI mass spectra of two mixtures of octanu­
cleotides (100 μM per strand) are shown in Figure 3.

As with the hexamers, the mass spectra show the

presence of noncovalent dimers (homodimers and het­
erodimers; see insets, Figure 3) in both the mixtures of

complementary and noncomplementary ODNs . The

ions that represent the heterodimer complexes in both

solutions are about 5% relative abundance. The ions

from the heterodimer in both mixtures appear to be

more abundant than might be predicted on the basis of

statistics. Table 1 summarizes the expected and ob­
served relative abundances for the hexamer and octa­
ter complexes. Although the relative abundances of

both the complementary and noncomplementary mix­
tures of the octamer do exceed that predicted values,

the observed relative abundance of the complementary

strands exceeds that predicted by a larger margin. One

must interpret the excess as an indication of some

degree of specificity in the interaction of the het­
erodimer that does not exist for the homodimers. (The

homodimers are assumed to represent nonspecific

binding.) Thus, both specific and nonspecific interac­
tions appear in the same spectrum. Close examination

of the data for the hexamer might also indicate a

degree of specificity of the complementarity het­
erodimer over the noncomplementary heterodimer, al­

though the effect is small. No duplex ions are observed

from solutions of octamer mixtures (either comple­

mentary or noncomplementary) with a concentration

of 10 μM for each strand.

The observation of any specificity in such short

ODNs is somewhat surprising. The predicted Tm for

the complementary strands is 7.4 °C for the octamer, so

at room temperature no stable solution interaction is

predicted, and there was no significant increase in

absorbance at 260 nm with increased temperature.

However, if one considers the potential for Watson–Crick base pairing of the most favorable

alignment of the two strands in each mixture, a rough

correlation can be made. Assuming a C-G base pair

contributes three hydrogen bonds and an A-T pair

contributes two hydrogen bonds, one can calculate the

total number of hydrogen bonds expected between

each pair of ODN strands. For example, d(TCTCTTTT)

and d(AGAGAAAA) are noncomplementary, but when

the 3' of one strand is aligned to the 5' end of the other

and the strands are offset by three bases, the sequence

TCTCT can base-pair to the sequence AGAGA. This

would provide three A-T pairs (two H bonds each)

and two C-G pairs (three H bonds each), for a total of

12 H bonds. Plotting the number of H bonds against

the excess of the measured relative abundance over

that predicted on a statistical basis gives a reasonably

linear correlation (r² = 0.972) as shown in Figure 4.

Although the number of hydrogen bonds is a primitive

measure of dimer stability and does not take into

Figure 3. Negative ion ESI mass spectra of binary mixtures of octamers at a concentration of 100 μM per strand of ODN in

10 mM aqueous ammonium bicarbonate at pH 8.5. (a) Mixture of complementary strands M1 = d(TCTTCTTTT) and M2 =
d(AGAGAAAAA). (b) Mixture of noncomplementary strands M1 =
d(TCTTCTTTT) and M2 = d(AGAGAAAAA). The insets are the

regions surrounding the mass-to-charge ratio of the triply charged
dimers, 2M1 and 2M2, represent homodimers; [M1 + M2] repre­
sents the heterodimer.

Figure 4. Correlation of excess mass spectral relative abundance with extent of predicted hydrogen bonding. The difference be­
tween the observed relative abundance and that predicted based

on statistics is plotted versus the number of hydrogen bonds in

the most favorable alignment of the ODN strands. Data from the

complementary and noncomplementary hexamers, octamers, and

decamers are included, as shown in Table 1.
account the effect of neighboring bases [21], most other common measures of duplex stability (e.g., $T_m$) are based on solution stability in the presence of solvent and salts. These solution-based measures may not be relevant to duplex stability in the gas phase. Our observation suggests that there is, in fact, a specific gas-phase interaction between ODN strands at temperatures well above those predicted for stable solution interaction. This may be due to the capture of very transient solution-phase complexes as gas-phase ions, and once formed in the gas phase in the absence of solvent, the duplexes have less tendency to dissociate than if they remained in solution.

If longer ODNs are studied, a predictable increase in specific interaction is noticeable. Figure 5 shows the ESI mass spectra of two mixtures of pentadecanucleotides: one pair is complementary and one pair is not. There is a striking increase in the relative abundance of the heterodimer in the complementary pair as compared to the noncomplementary pair. The predicted $T_m$ for the complementary duplex is 39.1 °C, so the stability of the duplex in solution is high enough for the complex to be observed at room temperature. The noncomplementary heterodimer also is present at a level that exceeds statistical prediction, consistent with the trend observed for the hexamers and octamers. A maximum of 27 hydrogen bonds might be expected if the noncomplementary sequences were aligned in the most favorable way, and there is a clear excess in the abundance of the heterodimer relative to the two homodimers. As further evidence of duplex stability in the case of pentadecamers, the mass spectrum of the complementary mixture at 10 μM per strand displays ions that correspond to the heterodimer, whereas the mass spectrum of the noncomplementary mixture does not.

It has been reported elsewhere that noncovalent interactions in ESI-MS can be disrupted by raising the energy of the incoming ions [8, 10–12]. Indeed, this has been used as a proof that these complexes are noncovalent [11, 12]. With the ODN complexes in this study, we see no evidence of this effect. Although the ions that represent the complexes can be removed by dilution of the solution, increasing the energy of the ion beam (increasing the orifice potential on the Sciex instrument) actually increases the abundance of the complexes, relative to the respective monomers. Figure 6 shows the effect of orifice potential on the absolute abundance of ions that represent the complexes. In nearly every case the ion abundance increases as the orifice potential increases in the negative direction, down to a value of −120 V. Below that value, the absolute abundance of the entire spectrum—monomer and dimer ions alike—drops precipitously. The increase in abundance of the noncovalent complex ions seems to contradict our understanding of the effect of ion energy in declustering complex ions, but the effect is real and reproducible. The oligonucleotides used in our studies were largely free of sodium, so sodium adducts did not contribute appreciably to the spectra and stripping of sodium adducts cannot explain the orifice effect. We have noted a similar effect in the ions that represent noncovalent complexes of a protein system and its associated peptide ligands (D. S. Wagner and R. J. Anderegg, unpublished results), and a recent literature report contains a similar observation for antisense peptide complexes [15].

Two other reports of mass spectrometry of oligonucleotide duplexes have appeared [12, 13]. The concen-
trations of ODNs used in those studies were at or above 100 \(\mu M\), well within the range where we see nonspecific complex formation.

**Conclusions**

Negative ion ESI mass spectra of oligonucleotides often show the presence of species with twice the molecular mass expected from the ODN. The appearance of these dimers is concentration dependent, which suggests that the association is noncovalent. When mixtures of complementary or noncomplementary ODNs are analyzed, homodimers and heterodimers are observed. The abundance of the dimers is statistically predicted unless a sufficient possibility for Watson–Crick base pairing exists, and there is a rough correlation between the extent of such pairing (hydrogen bonding) and the abundance of the dimer ions. Surprisingly, evidence of specific binding is observed in the mass spectra even at temperatures above the predicted \(T_m\), the maximum temperature at which stable solution dimers would be expected. Unlike previous studies, we find that the abundance of the dimers of the ODNs we tested seems to increase as the energy of the incoming ions increases until a potential where the overall abundance of the spectrum decreases significantly.

**Acknowledgments**

Helpful discussions with Steve Brown, Rich Bucholtz, Steve Haneline, and Jim Veal of Glaxo, Inc. are gratefully acknowledged. \(T_m\) measurements were made by Phil Myers of Glaxo.

**References**