

# Specific recognition of DNA bulge structure by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry

Yu-Ju Chen<sup>1\*</sup>, Po-Jui Chen<sup>1</sup>, Hsueh-Ting Wang<sup>2</sup> and Chien-Chung Cheng<sup>3</sup>

<sup>1</sup>Institute of Chemistry, Academia Sinica, Taipei, Taiwan 11529, ROC

<sup>2</sup>Department of Chemistry, National Changhua University of Education, Changhua Taiwan 500, ROC

<sup>3</sup>Department of Chemistry, Tamkang University, Taipei, Taiwan 251, ROC

Received 19 December 2003; Revised 8 January 2004; Accepted 11 January 2004

This study reports a novel approach utilizing an octahedral  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$  reagent in the presence of  $\text{H}_2\text{O}_2$  with analysis by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOFMS) to serve as an efficient probe for bulged DNA structures. Elucidation of DNA bulge-specific recognition pathways and cleavage mechanisms is demonstrated by characterization of bulge-specific cleavage products and other backbone lesion fragments. The cleavage specificity of  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2/\text{H}_2\text{O}_2$  arises from sugar oxidative strand scission, for which the position of the abstracted hydrogen is unambiguously determined as the 4'-H of the deoxyribose moiety. Furthermore, differentiation between bulge-specific recognition and diffusion-controlled non-selective cleavage can be clarified through time-dependent MALDI-TOFMS studies. The present results demonstrate that MALDI-TOFMS can be a sensitive and efficient technique for complex mechanistic studies of this kind, providing information for future rational drug design targeting bulged DNA structures. Copyright © 2004 John Wiley & Sons, Ltd.

Bulged structure in nucleic acid is one of the important motifs in DNA recognition. Its extra unpaired nucleobases are capable of forming complexes with nucleic acid-binding-proteins as well as acting as the binding site for small molecules.<sup>1–3</sup> They are believed to be the intermediates in the fragment shift mutagenesis inducing some genetic diseases.<sup>4–6</sup> Since DNA stores genetic information, a change in sequences or structures has the potential to induce erroneous gene expression resulting in the formation of non-functional proteins.

NMR spectroscopy and X-ray analysis are frequently used to reveal structural information. Despite its importance in biological recognition, only a few DNA bulge structures have been disclosed by X-ray crystallography or NMR methods due to the instability of such structures.<sup>7,8</sup> Recently, a novel  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$  complex was shown to serve as a bulge-specific probe, cleaving specifically at the bulge site of DNA.<sup>9</sup> The structure of this octahedral complex, deduced by X-ray crystallography, contains two fused 1,10-phenanthroline moieties, with all four pyridinium nitrogen atoms sitting on the same coordination plane, and two labile axial TFA ligands. The analysis of these radical-mediated cleavage products of nucleic acids has been a challenge, since the chemical similarities in the scission products make them difficult to

isolate or separate.<sup>10</sup> Cheng *et al.* used high-resolution polyacrylamide gel electrophoresis (PAGE) to visualize the scission products arising from the reaction of the  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$  complex with DNA.<sup>11</sup> However, due to the limited resolution of the gel electrophoresis, the precise cleavage position cannot be determined to elucidate the reaction mechanism. It is therefore desirable to develop a methodology capable of unambiguously identifying the cleavage products to enable deduction of the reaction mechanism.

The introduction of electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) as soft ionization methods has driven the direct and rapid identification of intact biomolecules. Historically, mass spectrometric experiments on oligonucleotides were performed in order to measure molecular weights of synthetic oligonucleotides.<sup>12</sup> With the invention of delayed-pulse ion extraction in time-of-flight mass spectrometry (TOFMS), the resolution of MALDI-TOFMS has been significantly increased. The resolution and mass accuracy of MALDI-TOFMS are sufficient to unambiguously identify reaction products. For instance, isomeric deoxyguanosine adducts were successfully differentiated based on their post-source decay (PSD) products using MALDI-TOFMS.<sup>13</sup> Compared with the complex spectra of multiply charged ions obtained using ESI, MALDI-TOFMS offers advantages of direct analysis of reaction mixture without prior separation and isolation. The RNase H reaction products resulting from digestion of oligoribonucleotide were directly analyzed by Polo and Limbach using nanomole amounts of material.<sup>12</sup>

\*Correspondence to: Y.-J. Chen, Institute of Chemistry, Academia Sinica, Taipei 11529, Taiwan, ROC.  
E-mail: yjchen@chem.sinica.edu.tw  
Contract/grant sponsors: Academia Sinica; National Science Council of the Republic of China.

As chemists are thinking about the rational design of new molecules for structure recognition on nucleic acids, it is obvious that the design must be based on a thorough understanding of the recognition mechanism. In this study we employed MALDI-TOFMS as a high-resolution tool to specifically detect DNA bulge structure and determine its sugar oxidative cleavage sites by  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$ . Complex DNA lesion fragments induced by  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$  can be characterized by MALDI-TOF mass spectra. In order to elucidate the strand scission pathways leading to specific DNA bulge recognition, the strategy employed in this study is to monitor the scission fragment pattern as a function of reaction time. We attempt to differentiate bulge-specific cleavage and non-specific cleavage pathways by their time-dependent competition. In addition to the specific bulged DNA recognition, the results also illustrate the potential of the MALDI-TOFMS technique in providing insights into complex mechanisms of nucleic acid damage.

## EXPERIMENTAL

### Materials

The synthetic 27-mer DNA substrate, HIV-27 (5'-GCA-GATCTGAGCCTGGGAGCTCTCTGC-3'), was purchased from GIBCO (Life Technologies, Taiwan) and purified by reversed-phase high-performance liquid chromatography (Agilent 1100 HPLC Systems, USA). The complex  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$  was prepared according to procedures described previously.<sup>14</sup> All reagents and ACS-grade solvents were purchased from commercial sources and were used without further purification.

### DNA cleavage by metal complexes

A total volume of 20  $\mu\text{L}$  solutions containing 20  $\mu\text{M}$  oligonucleotide and metal complex was mixed in the presence of 0.1%  $\text{H}_2\text{O}_2$  for varied times. The cleavage reaction was quenched by adding 3M  $\text{NH}_4\text{OAc}$  and 95% EtOH (60  $\mu\text{L}$ ), allowed to precipitate at  $-30^\circ\text{C}$  for 3 h, centrifuged (12 000 rpm) at  $4^\circ\text{C}$  for 45 min, and lyophilized to dryness to obtain a pellet. Enhanced cleavage efficiency was achieved by adding 60  $\mu\text{L}$  aqueous piperidine (0.7M) to the pellet and heating at  $90^\circ\text{C}$  for 15 min. The precipitated DNA and its fragments were lyophilized, washed with deionized water, and lyophilized again to dryness. The DNA sample was re-suspended with 5  $\mu\text{L}$   $\text{H}_2\text{O}$  to a final concentration of 80  $\mu\text{M}$ . 1  $\mu\text{L}$  of analyte-matrix solution containing 1  $\mu\text{L}$  DNA stock solution and 4  $\mu\text{L}$  matrix (50 mg/mL 3-hydroxypicolinic acid in a 1:1:2 mixture of water/acetonitrile/0.1 M ammonium citrate) was spotted onto the MALDI sample plate and air-dried.

### Mass spectrometry

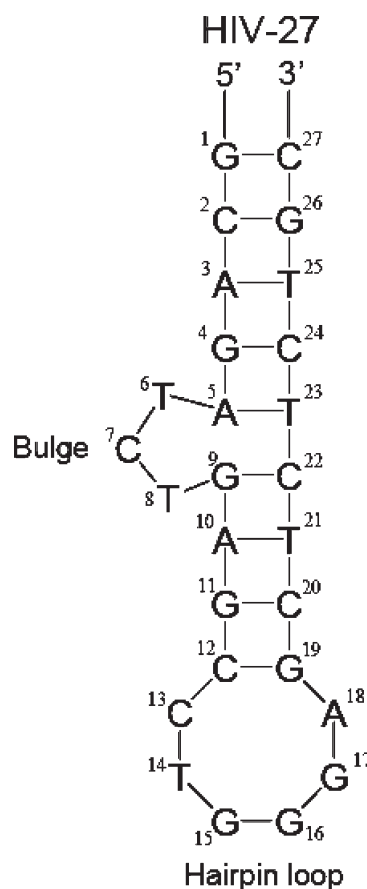
All experiments were performed using a Voyager DE-PRO mass spectrometer (Applied Biosystems, Framingham, MA, USA) equipped with a 337-nm nitrogen laser. The laser power was set just above the ionization threshold to avoid fragmentation of the oligonucleotides. Ions were accelerated through a potential of 25 kV. The mass spectrometer parameters optimized for the  $m/z$  range 1000–10 000 were as follows: delay time 300 ns, and guide wire potential 0.2% of the

accelerating voltage. To improve signal-to-noise ratio, typically 50 single-shot spectra were averaged. Internal calibration was performed with average  $m/z$  values for triply doubly and singly charged 27-mer HIV-27 of 2763.1, 4145.2 and 8291.4, respectively.

## RESULTS AND DISCUSSION

The choice of DNA sequence for the strand scission experiments was based on the RNA hairpin structure of the trans-activation response element (TAR-RNA).<sup>15,16</sup> A particular sequence of 27-mer DNA substrate, 5'-GCAGATCTGAGCCTGGGAGC-TCTCTGC-3', was chosen because its secondary structure comprises a three-base bulge, a six-base single-stranded loop, and double-stranded domains. The secondary structure of the 27-mer DNA is presented in Fig. 1. In this experiment, DNA was annealed prior to the cleavage reaction to ensure the presence of the hairpin structure.

Metal-cation adduction is a common problem for nucleic acid detection by mass spectrometry and various desalting techniques have been extensively discussed.<sup>17</sup> Considerable ion suppression caused by excess metal salt will reduce the sensitivity of mass spectrometric measurement. In addition, partial substitution of protons by metal cations will result in a distribution of adduct peaks and may interfere with



**Figure 1.** Secondary structure of HIV-27. The 27-mer DNA possesses bulge, hairpin loop and double-stranded regions. All the oligonucleotides are numbered for the convenience of assignment in the text.

interpretation of fragment peaks resulting from the DNA cleavage reaction. To reduce the amount of salt and avoid sample loss due to further desalting steps, annealing was performed in the presence of 100 mM ammonium acetate instead of the commonly used sodium chloride. Ammonium acetate can effectively replace metal ions attached to the phosphate backbone.<sup>17</sup> Unreactive free  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$  complex in solution was removed by ethanol precipitation after the reaction was terminated. Presence of excess  $\text{Co}^{\text{II}}(\text{HAPP})^{2+}$  greatly reduced the signal-to-noise ratio (data not shown).

The MALDI-TOF mass spectra in this study show almost exclusively molecular ions of the 27-mer DNA substrate or its fragments, and little metal-cation adduction was observed. Due to the reduced lifetime, extensive ion fragmentation of oligonucleotides becomes more significant with increasing masses in MALDI-TOFMS,<sup>17</sup> and will also interfere with interpretation of the cleavage products. Therefore, the MALDI-TOFMS parameters, especially the laser power, were optimized carefully to prevent artifactual DNA fragmentation.

### Characterization of HIV-27 cleavage products

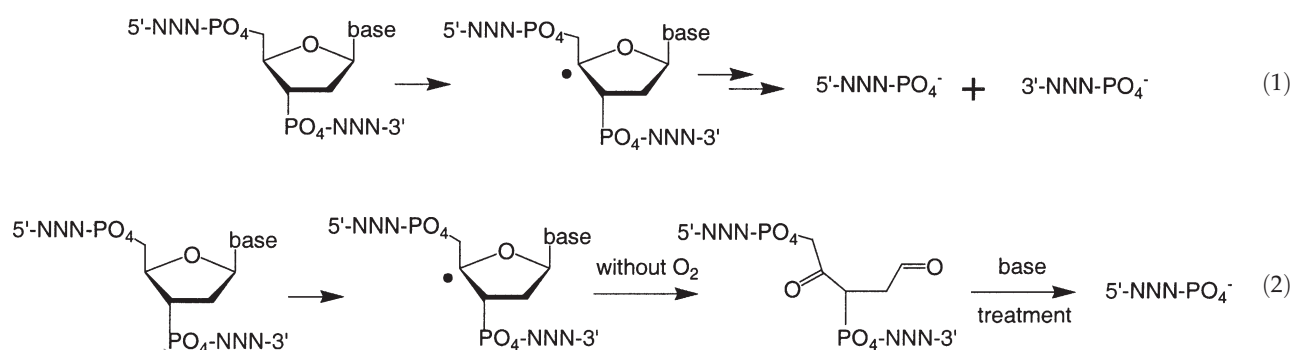
To accurately characterize the cleavage products, we list the assignments and the expected  $m/z$  values of possible cleavage products in Table 1. The predicted  $m/z$  values were calculated from the average molecular weights of the oligonucleotide sequences of predicted cleavage products. The mass accuracy is better than 1 Da after internal calibration with average  $m/z$  values for triply, doubly and singly charged 27-mer HIV-27 of 2763.1, 4145.2 and 8291.4, respectively. As shown in Fig. 2(a), the MALDI-TOF mass spectrum obtained 30 min after reaction displays several series of fragment ions which are roughly spaced by the molecular weight of an 'average' mononucleotide. The dominant scission products ( $m/z$  1278.1, 1591.6, 1895.9, 2185.3, 2489.8 and 2818.7) were attributed to cleavage at the 5'-NNN- $\text{PO}_4$ -C-3' site of the bulge position, yielding oligonucleotide phosphate fragments 5'-NNN- $\text{PO}_4$  in different lengths, where NNN=GCAG(<sup>5</sup>G4), GCAGA(<sup>5</sup>A5), GCAGAT(<sup>5</sup>T6), GCAGATC(<sup>5</sup>C7), GCA-GATCT(<sup>5</sup>T8) and GCAGATCTG(<sup>5</sup>G9), respectively. The relative fragment distribution agrees well with the gel electrophoresis result,<sup>9</sup> and clearly demonstrates that fragmentation resulting from cleavage at the bulge positions of

<sup>5</sup>A5, <sup>5</sup>T6, <sup>5</sup>C7 and <sup>5</sup>T8 is a major reaction channel. Cleavages at sites near the flanking junctions such as <sup>5</sup>G4 and <sup>5</sup>G9 were also observed, although the intensities of <sup>5</sup>G9 drop significantly. The guanine residues in this DNA loop have been reported to be most susceptible to oxidative cleavage due to their accessibility and low reduction potential.<sup>14</sup> However, no significant enhanced cleavage was observed in the 5'-TGGGA-3' region of the hairpin loop of HIV-27. Though collision-induced dissociation of oligonucleotides also generates backbone cleavage products,<sup>18,19</sup> it should be noted that the loss of neutral base, which usually accompanies the formation of 5'-sequence fragments, was not observed as a relevant fragmentation channel. This suggested that the mechanism of DNA cleavage by  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$  in solution is different from the fragmentation process by mass spectrometric dissociation in the gas phase.

In addition to the 5'-NNN- $\text{PO}_4$  fragments, cleavage at the 5'-C- $\text{PO}_4$ -NNN-3' site was also observed as a minor channel, resulting in the 3'-NNN- $\text{PO}_4$  fragments in Fig. 2(a). Unlike the 5'-NNN- $\text{PO}_4$  fragments, these 3'-NNN- $\text{PO}_4$  fragments have a wide  $m/z$  distribution from 1000 to about 7000 ( $m/z$  1229.5, 1533.5, 1823.0, 2127.4, 2417.0, 2746.1, 3059.7, 3388.7, 3717.7, 4045.8, 4350.3, 4639.6, 4929.0, 5257.6, 5571.4, 5900.4, 6204.5, 6493.4). The assignments shown in Table 1 indicate that these 3'-sequence fragments are from cleavages at both bulge (<sup>3</sup>C7, <sup>3</sup>T8, <sup>3</sup>G9 and <sup>3</sup>A10) and hairpin loop (<sup>3</sup>T14, <sup>3</sup>G15, <sup>3</sup>G16, <sup>3</sup>G17, <sup>3</sup>A18 and <sup>3</sup>G19) regions, and at other sites. Unexpectedly, some of these low-abundance 3'-sequence fragments were produced from cleavage at <sup>3</sup>C24, <sup>3</sup>T23, <sup>3</sup>C22, <sup>3</sup>T21 and <sup>3</sup>C20; these positions are opposite to the bulge structure rather than at the bulge sites.

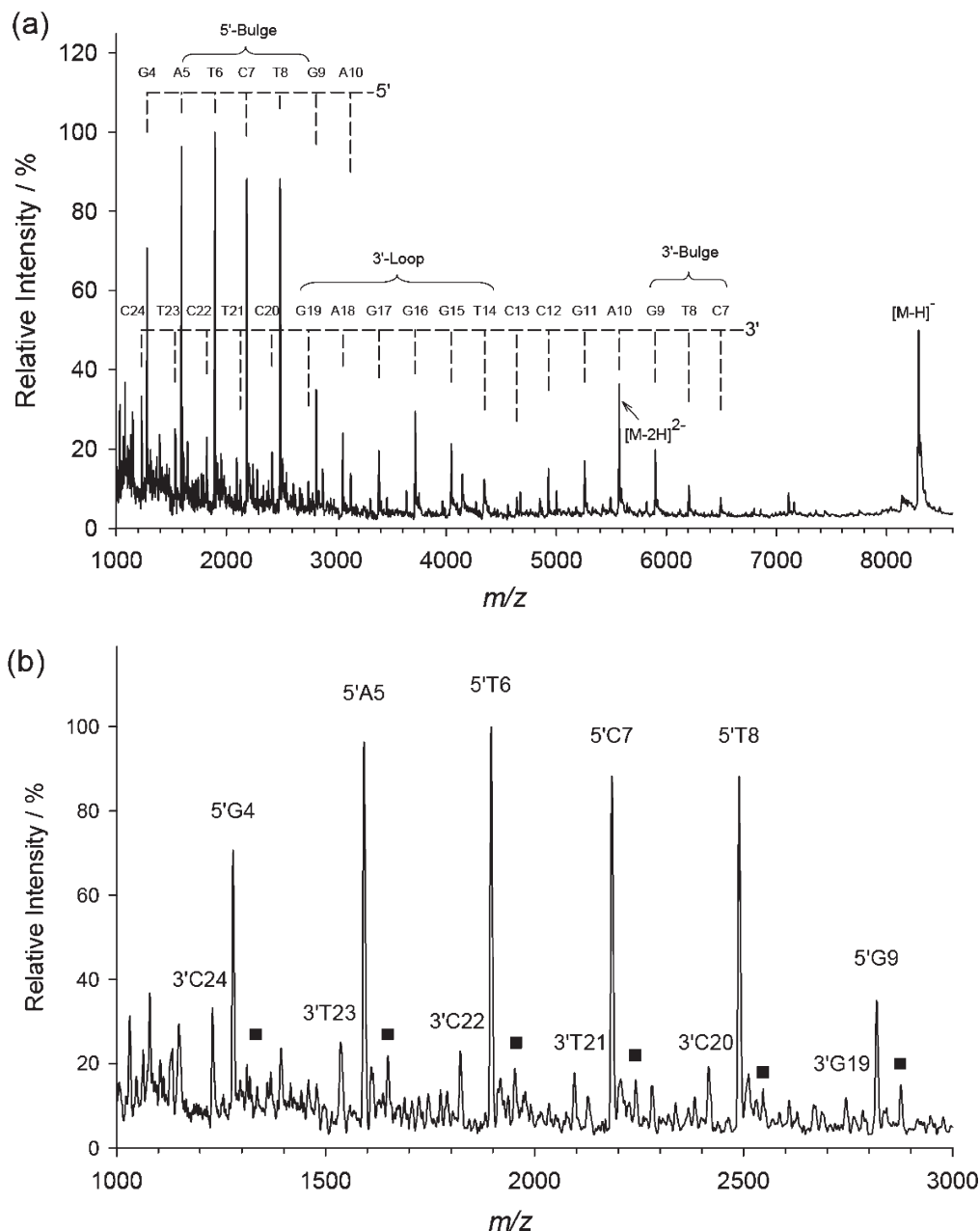
### DNA cleavage pathways by the $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$ complex

In attempting to precisely assess the mechanism of the DNA cleavage reaction, the position at which the deoxyribose moiety of DNA is oxidized during  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$ -catalyzed DNA cleavage has to be determined.<sup>21</sup> Characterization of the terminal structures of the DNA cleavage products can reveal the sugar oxidative position. Based on the above 5'-NNN- $\text{PO}_4$  and 3'-NNN- $\text{PO}_4$  fragments, we propose the cleavage mechanism by the  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$  complex shown in Eqns. (1) and (2).



**Table 1.** List of possible cleavage products and their  $m/z$  values for the reaction of  $\text{Co}^+(\text{HAPP})(\text{TFA})_2$  with 27-mer DNA. The expected  $m/z$  values of these 5'-NNN- $\text{PO}_4$  series or 3'-NNN- $\text{PO}_4$  series fragments correspond to average  $[\text{M}-\text{H}]^-$  values. Fragments observed in this experiment are underlined and in bold font

Label	Fragment sequence	Theory/Experiment $m/z$	Label	Fragment sequence	Theory/Experiment $m/z$
$^5\text{G1}$	<b>5'-GCAGATCTGAGCCTGGGAGCTCTCTGCP</b>	<b>8291.4</b>	$[\text{M}-\text{H}]^-$		
$^5\text{C2}$	5'-Gp		$^3\text{C2}$	pCAGATCTGAGCCTGGGAGCTCTCTGC-3'	8043.2
$^5\text{A3}$	5'-GCp	635.4	$^3\text{A3}$	pAGATCTGAGCCTGGGAGCTCTCTGC-3'	7754.0
$^5\text{G4}$	5'-GCAP	948.6	$^3\text{C4}$	pGATCTGAGCCTGGGAGCTCTCTGC-3'	7440.8
$^5\text{A5}$	<b>5'-GCAGp</b>	<b>1277.8/1278.1</b>	$^3\text{A5}$	pAICTGAGCCTGGGAGCTCTCTGC-3'	7111.6
$^5\text{T6}$	5'-GCAGAp	<u>1591.0/1591.6</u>	$^3\text{T6}$	pTCTGAGCCTGGGAGCTCTCTGC-3'	6798.4
$^5\text{T6}$	<b>5'-GCAGATp</b>	<b>1895.2/1895.9</b>	$^3\text{C7}$	<b>pCTGAGCCTGGGAGCTCTCTGC-3'</b>	<b>6494.2/6493.4</b>
$^5\text{C7}$	5'-GCAGATp	<u>2184.4/2185.3</u>	$^3\text{T8}$	<u>pTGAGCCTGGGAGCTCTCTGC-3'</u>	<u>6205.0/6204.5</u>
$^5\text{T8}$	<b>5'-GCAGATCP</b>	<b>2488.6/2489.8</b>	$^3\text{G9}$	<b>pAGCCTGGGAGCTCTCTGC-3'</b>	<b>5900.8/5900.4</b>
$^5\text{G9}$	5'-GCAGATCP	<u>2817.8/2818.7</u>	$^3\text{A10}$	<u>pAGCCTGGGAGCTCTCTGC-3'</u>	<u>5571.6/5571.4</u>
$^5\text{A10}$	5'-GCAGATCTGAp	3131.0	$^3\text{G11}$	<u>pGCCCTGGGAGCTCTCTGC-3'</u>	<u>5258.4/5257.6</u>
$^5\text{G11}$	5'-GCAGATCTGAp	3460.4	$^3\text{C12}$	<b>pCCTGGGAGCTCTCTGC-3'</b>	<b>4929.2/4929.0</b>
$^5\text{C12}$	5'-GCAGATCTGAGp	3749.4	$^3\text{C13}$	pCTGGGAGCTCTCTGC-3'	4640.0/4639.6
$^5\text{C13}$	5'-GCAGATCTGAGCp	4038.6	$^3\text{T14}$	<u>pTGGGAGCTCTCTGC-3'</u>	<u>4350.8/4350.3</u>
$^5\text{T14}$	<b>5'-GCAGATCTGAGCCTp</b>	<b>4342.8/4342.2</b>	$^3\text{G15}$	<b>pGGGAGCTCTCTGC-3'</b>	<b>4046.6/4045.8</b>
$^5\text{G15}$	5'-GCAGATCTGAGCCTGp	<u>4672.0/4672.1</u>	$^3\text{G16}$	<u>pGGGAGCTCTCTGC-3'</u>	<u>3717.9/3717.7</u>
$^5\text{G16}$	<b>5'-GCAGATCTGAGCCTGGp</b>	<b>5001.2/5001.0</b>	$^3\text{G17}$	<b>pGAGCTCTCTGC-3'</b>	<b>3388.2/3388.7</b>
$^5\text{G17}$	5'-GCAGATCTGAGCCTGGGp	5330.4	$^3\text{A18}$	<u>pAGCTCTCTGC-3'</u>	<u>3059.0/3059.7</u>
$^5\text{A18}$	5'-GCAGATCTGAGCCTGGGAp	5643.7	$^3\text{G19}$	<u>pGCCCTCTGC-3'</u>	<u>2745.8/2746.1</u>
$^5\text{G19}$	5'-GCAGATCTGAGCCTGGGAGp	5672.9	$^3\text{C20}$	<u>pCTCTCTGC-3'</u>	<u>2416.6/2417.0</u>
$^5\text{C20}$	5'-GCAGATCTGAGCCTGGGAGCp	6262.0	$^3\text{T21}$	<u>pTCTCTGC-3'</u>	<u>2127.4/2127.4</u>
$^5\text{T21}$	<b>5'-GCAGATCTGAGCCTGGGAGCTp</b>	<b>6566.2</b>	$^3\text{C22}$	<b>pCTCTGC-3'</b>	<b>1823.2/1823.0</b>
$^5\text{C22}$	5'-GCAGATCTGAGCCTGGGAGCTp	6855.4	$^3\text{T23}$	<u>pTCTCTGC-3'</u>	<u>1534.0/1533.5</u>
$^5\text{T23}$	5'-GCAGATCTGAGCCTGGGAGCTCp	7159.6	$^3\text{C24}$	<b>pCTGC-3'</b>	<b>1229.8/1229.5</b>
$^5\text{C24}$	5'-GCAGATCTGAGCCTGGGAGCTCp	7448.8	$^3\text{T25}$	<u>pTGC-3'</u>	<u>940.6</u>
$^5\text{T25}$	5'-GCAGATCTGAGCCTGGGAGCTCTp	7753.0	$^3\text{C26}$	<u>pCC-3'</u>	<u>636.4</u>
$^5\text{C26}$	5'-GCAGATCTGAGCCTGGGAGCTCTCTp	8082.2	$^3\text{C27}$	<u>pC-3'</u>	<u>307.2</u>

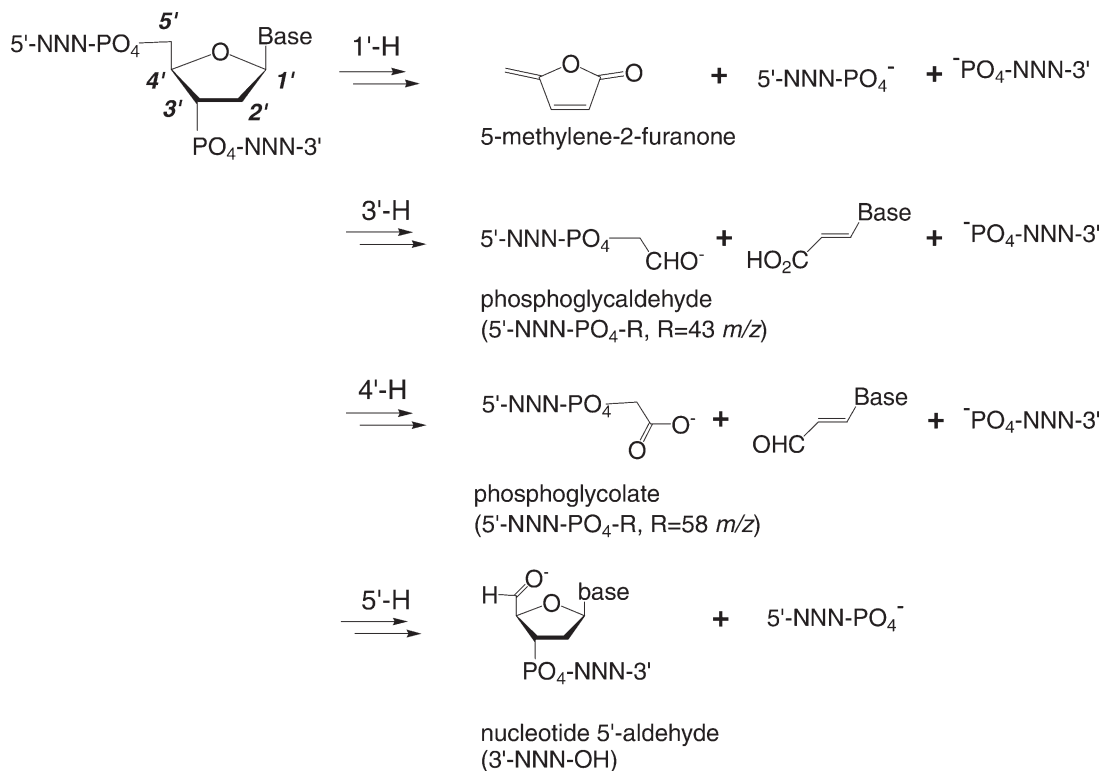


**Figure 2.** (a) MALDI-TOF mass spectrum of products from the reaction of  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$  with 27-mer DNA in 0.1%  $\text{H}_2\text{O}_2$  for 30 min. Assignment of these 5'-NNN- $\text{PO}_4$  and 3'-NNN- $\text{PO}_4$  cleavage fragments are listed in Table 1. Ions produced from a minor cleavage channel, whose molecular weights are 58 Da higher than the 5'-NNN- $\text{PO}_4$  fragments, are marked with filled squares (■). The spectra in the  $m/z$  range 1000–3000 is expanded in (b).

Nucleotide 3'-phosphate (5'-NNN- $\text{PO}_4$ ) and nucleotide 5'-phosphate (3'-NNN- $\text{PO}_4$ ) are common DNA damage products from agents including activated antibiotics, metal complexes and molecules ionized by high-energy radiation.<sup>21</sup> The formation of these sugar-phosphate backbone cleavage products is believed to be initiated by hydrogen abstraction. A key question in elucidating DNA cleavage mechanism is: which hydrogen atom in a sugar moiety is abstracted to activate the strand scission. Another minor product formation channel was also observed (and is zoomed) in Fig. 2(b), producing ions at  $m/z$  1649.0, 1952.8, 2242.2, 2546.3 and 2875.6 (peaks marked with filled squares ■). The  $m/z$  values for these cleavage products are greater than those of dominant bulge-

specific 5'-NNN- $\text{PO}_4$  fragments by 58 Th (NNN = GCAG, GCAGA, GCAGAT, GCAGATC, GCAGATCT and GCA-GATCTG), suggesting the possibility of a reaction channel other than those shown in Eqns. (1) and (2). The sugar oxidation will generate characteristic fragments in the reaction mixture. The different terminal structures of these fragments (Scheme 1) could reveal information about the mechanism of deoxyribose oxidation.<sup>20</sup>

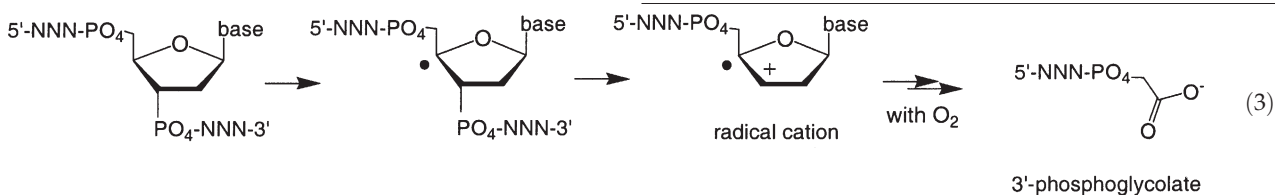
For instance, neither phosphoglycaldehyde (5'-NNN- $\text{PO}_4$ -R, R = 43 Da) nor nucleotide 5'-aldehyde (3'-NNN-OH) was observed during  $\text{CoHAPP}^{2+}/\text{H}_2\text{O}_2$  catalyzed DNA cleavage, indicating the cleavage mechanism does not proceed via the 3'-H or 5'-H abstraction pathways. The increase of 58 Da units is



**Scheme 1.** Sugar oxidation at different positions generates characteristic fragments with different terminal structures.

most likely to be the 3'-terminal phosphoglycoate ( $5\text{-NNN-PO}_4\text{-CH}_2\text{CO}_2^-$ ). Based on the unique marker products, [ $5\text{-NNN-PO}_4+58$ ] fragments, we propose that the 4'-H abstraction pathway is responsible for DNA bulged cleavage shown in Eqn. (3). This series of marker fragments arises specifically from cleavage at positions starting from  $^5\text{G4}$ ,  $^5\text{A5}$ ,  $^5\text{T6}$ ,  $^5\text{C7}$ ,  $^5\text{T8}$ , and  $^5\text{G9}$ , which strongly suggests that the 4'-H abstraction mechanism may be responsible for the HIV-27 bulge recognition.

whether these products are simply kinetically stable intermediates from competitive reaction channels or are thermodynamically stable products. To further elucidate the temporal correlation of different reaction schemes, the MALDI-TOFMS analysis was performed at different reaction times. The variation of relative intensities of different cleavage fragments was monitored as a function of time. Figure 3 shows the MALDI-TOF mass spectra obtained at 5,



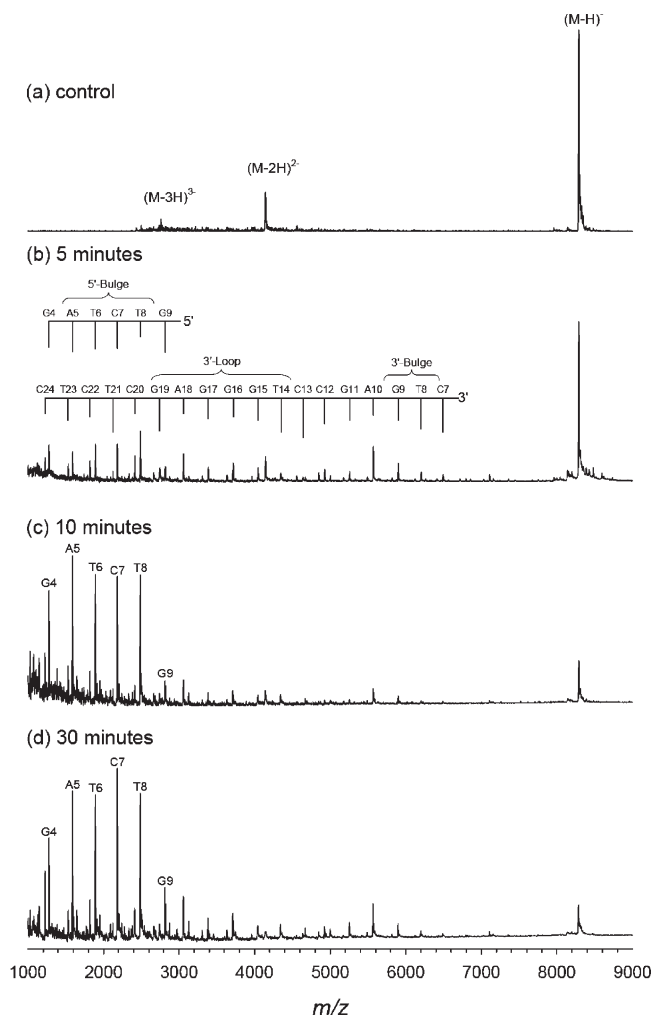
The conventional way to analyze DNA cleavage products involves a tedious procedure and requires large amounts of sample. Without detecting the intermediate radical or neutral products in the reaction, we can precisely locate the position of the abstracted hydrogen responsible for specific bulged-DNA cleavage. The MALDI-TOFMS analysis offers a convenient and accurate way to directly characterize the DNA fragment terminus, and thus to further elucidate different reaction pathways.

### Time-dependent study on bulge-specific recognition

Thus far we have demonstrated bulge-specific recognition by  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$  and deduced three strand scission pathways for the cleavage products. Without any information on the reaction kinetics, we cannot answer the question

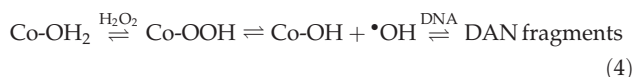
10 and 30 min, along with a control experiment. The control experiment was performed under the same reaction conditions but without  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$ . From the dramatically different fragmentation patterns at various times, we may expect that the strand scission of the 27-mer DNA is controlled by several competing processes. Thus, the key to understanding oxidative DNA scission relies not only on the cleavage mechanistic pathways by which  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$  cleaves DNA, but also on the kinetic aspects of DNA cleavage profiles. Based on several noteworthy observations from the time-dependent experiments, we will use Eqns. (1), (2) and (3) to postulate a hypothesis for the bulge-specific cleavage that follows.

Five minutes after the reaction had occurred, the complicated MALDI-TOF mass spectrum of the reaction mixture shown in Fig. 3(b) was obtained. Both the  $5\text{-NNN-PO}_4$  and



**Figure 3.** MALDI-TOF mass spectra from reaction of  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$  with 27-mer DNA for (a) control experiment (no  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$ ), and after reaction times of (b) 5 min, (c) 10 min, and (d) 30 min. Top annotations on fragments labeled 5' in (b) correspond to 5'-NNN- $\text{PO}_4$  series ions. Bottom assignments on fragments labeled 3' correspond to 3'-NNN-

3'-NNN- $\text{PO}_4$  series of fragments, covering the wide  $m/z$  range from 1000 to 7000, were observed, whereas [5'-NNN- $\text{PO}_4$ +58] products arising from 4'-H sugar oxidation have not reached a detectable level at this early stage. This implies that the sugar-phosphate backbone is highly vulnerable to attack by the  $\text{Co}^{\text{II}}(\text{HAPP})^{2+}$ , resulting in lesions in the 5'- or 3'-phosphodiester bond, as depicted in Eqns. (1) and (2). The strand scission patterns have not yet revealed specific cleavage at the bulge position, but have rather shown extensive breakage at the bulge ( $^5\text{A5-G9}$ ,  $^3\text{C7-G9}$ ), hairpin loop ( $^3\text{C13-G19}$ ), flanking junctions ( $^5\text{G4}$ ,  $^3\text{A10-C12}$ ) and even the double-stranded region ( $^3\text{C20-C24}$ ). Generally, many metal complexes that cause oxidative strand scission of nucleic acids will generate hydroxyl radicals in the presence of  $\text{H}_2\text{O}_2$ . Similarly, the hydroxyl radical will form and be involved in  $\text{CoHAPP}^{2+}/\text{H}_2\text{O}_2$  catalyzed DNA cleavage,<sup>11</sup> as illustrated in Eqn. (4).



In the initial stage of reaction, attraction between DNA and metal complexes in solution is mainly electrostatic. When the positively charged  $\text{Co}^{\text{II}}(\text{HAPP})^{2+}$  is attracted to the polyanionic backbone of the DNA by electrostatic interactions, the produced neutral hydroxyl radical then becomes diffusive and highly reactive toward different sites of the DNA. Due to the lack of specificity, this kind of oxidative cleavage will not show a unique fragmentation pattern. Observation of a wide distribution of both the 5'-NNN- $\text{PO}_4$  and 3'-NNN- $\text{PO}_4$  fragments at 5 min is indicative of the diffusive nature of DNA cleavage via Eqn. (1).

If the diffusive hydroxyl radical is the only controlling factor responsible for the strand scission, one would expect fragments from different cleavage sites irrespective of the length of the reaction period. Under this assumption, we should observe similar fragment profiles throughout the entire reaction. At 10 min, however, a more simple fragment pattern predominantly in the  $m/z$  range 1000–3000 (Fig. 3(c)) was observed, and assigned to products from only the bulge position ( $^5\text{A5}$ ,  $^5\text{T6}$ ,  $^5\text{C7}$ ,  $^5\text{T8}$  and  $^5\text{G9}$ ). Although fragment  $^5\text{A10}$  from cleavage at the flanking junctions was also observed, its intensity is weaker compared with that of the bulge cleavage products. The bulge-specific cleavage pattern was not obviously in agreement with the predicted 3'-NNN- $\text{PO}_4$  fragments, and, therefore, Eqn. (2) is more likely to be responsible for the dominant bulge recognition fragments. No specific cleavage resulting in 5'-NNN- $\text{PO}_4$  fragments at this DNA hairpin loop was observed, despite the fact that the hairpin loop contains the same 5'-CTG-3' sequence as the bulge site. This observation appears to eliminate the possibility that the DNA bulge-specific recognition is due to sequence selectivity. Minor sugar oxidation products from the bulge position corresponding to Eqn. (3), i.e., 5'-NNN- $\text{PO}_4$ - $\text{CH}_2\text{CO}_2^-$  fragments at  $^5\text{A5}$ ,  $^5\text{T6}$ ,  $^5\text{C7}$ ,  $^5\text{T8}$  and  $^5\text{G9}$ , start to be observed at 10 min. The appearance of this characteristic fragment pattern clearly distinguishes the structural selectivity of the mechanism illustrated in Eqn. (3) between bulge, hairpin loop and single-stranded DNA regions. The late appearance of the specific bulge cleavage pattern at 10 min indicates that bulge-specific recognition via Eqns. (2) and (3) are slower reactions than the diffusion-controlled DNA cleavage mechanism shown in Eqn. (1).

Based on the present study for elucidation of the cleavage mechanism, the presence of both series of fragments, 5'-NNN- $\text{PO}_4$  and 5'-NNN- $\text{PO}_4$ - $\text{CH}_2\text{CO}_2^-$ , can be used as indicators for DNA bulge recognition. The specific cleavage patterns at all bulge positions will facilitate prediction for the bulge position of DNA with known sequences.

How does the  $\text{Co}^{\text{II}}(\text{HAPP})^{2+}$  species specifically recognize and cleave bulge sites? The observation that these specific bulge cleavage products replaced the non-specific fragmentation products at 10 min may suggest that the bulge-specific recognition mechanism substitutes for the pathway involving diffusive hydroxyl radical attack. The MALDI-TOF mass spectra reflect the transition from non-specific strand scission at 5 min to bulge-specific recognition at 10 min.

The solution structure of a two-base DNA bulge has been reported by NMR analysis. This structure was found to be a triangular prism pocket that binds with the enediyne analogs.<sup>21</sup> Though the non-covalent DNA- $\text{Co}^{\text{II}}(\text{HAPP})^{2+}$

complex was not observed by MALDI-TOFMS, this interaction is likely via the ligands and the bulge region in the DNA. Unlike bulge structures, the pocket of a hairpin loop is commonly full of nucleobases to form an unusual base-pairing conformation. Thus, there may not be enough space for a hairpin loop to host these  $\text{Co}^{\text{II}}(\text{HAPP})^{2+}$  ions, if no excess  $\text{Co}^{\text{II}}(\text{HAPP})^{2+}$  is present in the reaction.

Thirty minutes after the reaction, the 5'-NNN- $\text{PO}_4$  scission fragments at the bulge position are still the dominant species (Fig. 3(d)), indicating that bulge recognition is the preferred and probably thermodynamically controlled reaction pathway. Surprisingly, products of other minor reaction channels, such as 3'-NNN-5'- $\text{PO}_4$  that involve cleavage at the loop and especially at G17, G16 and G15, start to be apparent at this reaction time. This could be explained by the late occurrence of secondary reaction channels. At 30 min, after most of the bulge sites have been occupied by  $\text{Co}^{\text{II}}(\text{HAPP})^{2+}$ , excess  $\text{Co}^{\text{II}}(\text{HAPP})^{2+}$  may start to bind to hairpin loop and even double-stranded regions in the DNA. The subsequent cleavage reactions became non-specific and yield complex strand scission patterns. As a consequence,  $^3\text{T}14$ ,  $^3\text{G}15$ ,  $^3\text{G}16$ ,  $^3\text{G}17$  and  $^3\text{A}18$  show relatively higher intensities than those observed at 10 min, although bulge recognition is still the major reaction pathway.

## CONCLUSIONS

The MALDI-TOFMS technique has been used to successfully demonstrate the bulge-specific recognition of DNA by the octahedral complex  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$ . Compared with classical electrophoretic methods, the complex fragment products resulting from DNA strand scission can be unambiguously assigned using accurate molecular weight determinations. Three strand scission pathways have been elucidated based on the various fragment ion series. The sugar-phosphate backbone is most vulnerable to attack by the  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$  complex, producing lesion products consisting of predominantly 3'-phosphate termini (5'-NNN- $\text{PO}_4$ ) fragments and significantly smaller amounts of 5'-phosphate termini(3'-NNN- $\text{PO}_4$ ) fragments. The observation of 3'-phosphoglycolate (5'-NNN- $\text{PO}_4\text{-CH}_2\text{CO}_2$ ) serves as a unique marker that helps to locate the position of hydrogen abstraction for specific bulged DNA recognition. The oxidative cleavage in ribose sugar moieties induced by  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$  is initiated by hydroxyl radical attacking at the 4'-H position, leading to bulge-specific cleavage. In addition, competitions between bulge-specific and non-selective strand scission events can be clarified through time-dependent MALDI-TOFMS studies. By monitoring the lesion products at different reaction times, thermodynamically

stable bulge recognition products dominate over diffusion-controlled non-selective cleavage after 10 min.

This study provides the first attempt to utilize combination of a novel  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$  reagent and MALDI-TOFMS analysis to show specific targeting and direct cleavage at the DNA bulge structure. The unprecedented accuracy and sensitivity of MALDI-TOFMS have great potential to serve as an efficient bulge structure probe, even for unknown sequences at bulge sites. Since bulges and internal loops are the most common but important branched structures in nucleic acids, particularly in RNA, this study may provide significant and preliminary information in the study of the relationship of the complicated RNA structures and their biological functions.

## Acknowledgements

The authors thank Prof. Wen Jwu Wang (Tamkang University) for kindly providing the  $\text{Co}^{\text{II}}(\text{HAPP})(\text{TFA})_2$  complex and Prof. Chau-Chung Han (Academia Sinica) for helpful suggestions and review of the manuscript. Financial support from Academia Sinica and the National Science Council of the Republic of China is gratefully acknowledged.

## REFERENCES

1. Turner DH. *Curr. Opin. Struct. Biol.* 1992; **2**: 334.
2. Lilley DMJ. *Proc. Natl Acad. Sci. USA* 1995; **92**: 7140.
3. Wang Y-H, Bortner CD, Griffith J. *J. Biol. Chem.* 1993; **268**: 17571.
4. Streisinger G, Okada Y, Emrich J, Newton J, Tsugita A, Terzaghi E, Inouye M. *Cold Spring Harbor Symp. Quant. Biol.* 1996; **31**: 77.
5. Streisinger G, Owen J. *Genetics* 1985; **109**: 633.
6. Kunkel TA. *Biochemistry* 1990; **29**: 8003.
7. Zhu J, Wartell RM. *Biochemistry* 1999; **38**: 15986.
8. Ke S-H, Wartell RM. *Biochemistry* 1995; **34**: 4593.
9. Cheng CC, Kuo Y-N, Chuang K-S, Luo C-F, Wang WJ. *Angew. Chem. Int. Edn. Engl.* 1999; **38**: 1255.
10. Pogożelski WK, Tullius TD. *Chem. Rev.* 1998; **98**: 1089.
11. Cheng CC, Huang-Fu WC, Hung KC, Chen PJ, Wang WJ, Chen YJ. *Nucleic Acids Res.* 2003; **31**: 2227.
12. Polo LM, Limbach PA. *J. Mass Spectrom.* 1998; **33**: 1226.
13. Chiarelli MP, Wu H-P, Autunes AM, Branco PS. *Rapid Commun. Mass Spectrom.* 1999; **13**: 2004.
14. Chang TH, Ong CW, Chou YM, Chuang KS, Wang WJ. *J. Chin. Chem. Soc.* 1996; **43**: 73.
15. Carter PJ, Cheng C-C, Thorp HH. *J. Am. Chem. Soc.* 1998; **120**: 632.
16. Neenhold HR, Rana TM. *Biochemistry* 1995; **34**: 6303.
17. Nordhoff E, Kirpekar F, Roepstorff P. *Mass Spectrom. Rev.* 1996; **15**: 67.
18. Grotjahn L, Blocker H, Frank R. *Biomed. Mass Spectrom.* 1985; **12**: 514.
19. McLuckey SA, Vaidyanathan, Habibi-Goudarzi GS. *J. Mass Spectrom.* 1995; **30**: 1222.
20. Pogożelski WK, Tullius TD. *Chem. Rev.* 1998; **98**: 1089.
21. Maxam AM, Gilbert W. *Proc. Natl Acad. Sci. USA* 1997; **74**: 560.