3866–3872 *Nucleic Acids Research, 1996, Vol. 24, No. 19*

Characterization of peptide–oligonucleotide heteroconjugates by mass spectrometry

Ole N. Jensen[+], Sandhya Kulkarni[1,§], Jane V. Aldrich[1,¶] and Douglas F. Barofsky[2,*]

Department of Biochemistry and Biophysics, [1]College of Pharmacy, [2]Department of Agricultural Chemistry and Environmental Health Sciences Center, Oregon State University, Corvallis, OR 97331-7301, USA

Received April 8, 1996; Revised and Accepted August 16, 1996

ABSTRACT

Two peptide–oligothymidylic acids, prepared by joining an 11 residue synthetic peptide containing one internal carboxyl group (Asp side chain) to amino-linker-5′pdT_6 and amino-linker-5′pdT_{10} oligonucleotides, were analyzed by matrix-assisted laser desorption/ionization (MALDI) on a linear time-of-flight mass spectrometer and by electrospray ionization (ESI) on a triple–quadrupole system. These synthetic compounds model peptide–nucleic acid heteroconjugates encountered in antisense research and in studies that use photochemical crosslinking to investigate molecular aspects of protein–nucleic acid interactions. MALDI and ESI sensitivities for the two hybrid compounds were found to be similar respectively to their sensitivities for the pure oligonucleotide parts. In general, MALDI proved to be less affected by sample impurities and more sensitive than ESI, while ESI on the quadrupole produced greater mass accuracy and resolution than MALDI on the time-of-flight instrument. A hybrid's behavior in a MALDI-matrix or an ESI-spray-solvent was found to be governed mainly by the oligonucleotide. A single positive ESI tandem mass spectrum of the peptide-dT_6 accounted for the heteroconjugate's entire primary structure including the point of the oligonucleotide's covalent attachment to the peptide.

INTRODUCTION

Covalent protein–nucleic acid conjugates occur in nature, for example, in the genome of certain RNA and DNA viruses (1), in DNA topoisomerases (2), or as amino acid–RNA conjugates in aminoacyl-tRNAs (3). Covalent nucleoprotein complexes can be produced in the laboratory by irradiating cells or protein–nucleic acid mixtures with ultraviolet (UV) light (4–6). The bonds induced in such a photochemical reaction form at or near points of contact between protein and nucleic acid; thus, identification of the amino acids that are crosslinked to nucleobases can provide information about the nucleic acid binding domain of the protein (7). Peptide–oligonucleotide heteroconjugates have recently been synthesized for use in antisense research; the peptidic components of these hybrid complexes are thought to help stabilize, transport or target antisense oligonucleotides (8–13).

Recently, Jensen *et al.* described a protocol for studying protein–nucleic acid interactions (14,15) that makes extensive use of matrix-assisted laser desorption/ionization (MALDI) (16,17) and electrospray ionization (ESI) (18,19) mass spectrometric techniques to analyze photochemically crosslinked protein–nucleic acid complexes. Bioanalytical applications of MALDI and ESI techniques have been reviewed (20–22). MALDI and ESI mass analyses of a covalent peptide–oligonucleotide hybrid pose special problems because the two polymeric components generally have conflicting requirements for ionization. Some of these difficulties were overcome and certain others were delineated in recent investigations of various UV-crosslinked protein–nucleic acid systems (14,15,23).

In order to further develop mass analysis as a tool for investigating covalent nucleopeptides, we have chemically synthesized peptide–oligothymidylic acid conjugates and examined their behavior under MALDI and ESI mass spectrometric conditions. We report our observations and discuss their analytical import in this paper.

MATERIALS AND METHODS

Synthesis of peptide–oligonucleotide hybrids

The peptide was assembled on a PAL® resin (Millipore, 0.5 g) using amino acids protected with 9-fluorenylmethoxycarbonyl and a Biosearch 9500 automated peptide synthesizer as previously described (24), except that *N*,*N*-dimethylacetamide (HPLC grade, Aldrich) was used in place of *N*,*N*-dimethylformamide (DMF). The side chains of Arg and Asp were protected respectively with the groups 2,2,5,7,8-pentamethylchroman-6-sulfonyl and t-butyl. After assembly and removal of the 9-fluorenylmethoxycarbonyl group off the N-terminus, the peptide was treated with a 4-fold excess of acetyl 1-imidazole (0.4 M) for 2 h. The resin was washed and dried; then the peptide was released by treating the resin with trifluoroacetic acid (TFA) containing scavengers, Reagent K (25), for 3 h and isolated by a previously described method (24).

The peptide was purified on a Beckman Model 100A HPLC system equipped with a 4.6 mm × 250 mm, 5 µm particle, 300 Å pore, Vydac C_{18} column (The Separations Group) and a Waters Model 220 UV-detector. The mobile phase comprised (A) 0.1% TFA in water and (B) 0.08% TFA in acetonitrile (Burdick and Jackson); solvent was delivered at 1 ml/min with a gradient of 5–35% B for the first 10 min, 35–65% B for the next 30 min, and

* To whom correspondence should be addressed

Present addresses: [+]European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69012 Heidelberg, Germany, [§]Molecular Probes, 4849 Pitchford Avenue, Eugene, OR 97402, USA and [¶]University of Maryland at Baltimore, Department of Pharmaceutical Sciences, 20 North Pine Street, Baltimore, MD 21201-1180, USA

60–90% B for the last 5 min. Detection was at 220 nm. The peptide eluted at ~33.5 min.

The purified peptide was coupled to the aminolinker-oligodeoxythymidylic acids NH_2-$(CH_2)_6$-5′-p$(dT)_6$ and NH_2-$(CH_2)_6$-5′-p$(dT)_{10}$ (aminolinker-dT_6 and aminolinker-dT_{10} respectively; gifts from Oligos Etc., Inc., Wilsonville, Oregon) in separate reactions carried out in DMF (Burdick and Jackson) at room temperature. The DMF stock solution was sparged with N_2 for 2 h prior to use. HPLC-purified peptide (1.57 mg, 1.16 µmol) in 100 µl DMF was sonicated briefly in a 1.5 ml microcentrifuge tube. *N*,*N*-Diisopropylethylamine (Millipore; 10% in DMF, 8.5 µl), 1,2-benzotriazolyloxy-tris[pyrrolidino]phosphonium hexafluorophosphate (PyBOP; NovaBiochem; 100 mg/ml in DMF, 11.8µl), and 1-hydroxybenzotriazole (50 mg/ml in DMF, 7 µl) respectively were added to the peptide solution, and the mixture was allowed to react for 10 min. The peptide cocktail was then transferred to another microcentrifuge vial containing 820 µg of aminolinker-oligothymidylic acid dissolved in 73 µl DMF. The reaction solution was vortexed at room temperature for incubation periods that varied from 4 h to 3 days. Samples of the reaction mixture were analyzed at various intervals by MALDI mass spectrometry and microbore reverse-phase HPLC. The chromatography system consisted of a Perkin-Elmer ABI 140B syringe pump, a 1 mm × 250 mm, 5 µm particle, 300 Å pore, Vydac C_{18} column (The Separations Group), and an ABI/Kratos Model 783 UV-detector equipped with a microbore flow cell (1 µl dead volume). The solvent system was (A) 5 mM ammonium acetate and (B) acetonitrile; solvent was delivered at 40 µl/min with a gradient of 5–90% B over 45 min. The eluted species were monitored at 256 nm. The retention times of individually chromatographed synthetic peptide, aminolinker-dT_6, aminolinker-dT_{10} and PyBOP were used to locate the peaks corresponding to these primary reagents in the chromatograms of the reaction mixtures.

The crude peptide-linker–oligothymidylic acids were purified on the same HPLC system used to purify the synthetic peptide, but the solvents and gradient employed were those used on the microbore HPLC system to monitor the coupling reaction (see preceding paragraph). The flowrate was 1 ml/min, and the heteroconjugates were monitored at 256 nm.

Matrix-assisted laser desorption/ionization mass spectrometry

MALDI mass analyses were performed on a custom-built, linear time-of-flight mass spectrometer. Both the instrument and the operating conditions have been described elsewhere (15). Each mass spectrum was produced from 30 consecutive pulses of 355 nm photons from a Nd:YAG laser (Spectra Physics GCR-11). Signals from matrix ions or ions from various standards added to the sample were used for mass calibration.

Prior to being mass analyzed, aliquots from the reaction mixtures were diluted 100-fold with Millipore-filtered water, and HPLC purified samples were lyophilized and redissolved in 50–200 µl water/acetonitrile (1:1, v/v). A saturated solution of 4-hydroxy-α-cyanocinnamic acid (Aldrich) in 0.1% TFA/33% acetonitrile (2:1, v/v) (26) was used as the laser desorption/ionization matrix for the synthetic peptide, and a saturated solution of 2,4,6-trihydroxyacetophenone (Aldrich) in 50 mM diammonium hydrogen citrate/50% acetonitrile (27) was used as the matrix for the aminolinker–oligothymidylic acids, the crude reaction mixtures, and the HPLC-purified peptide-linker–oligothymidylic acids. Before depositing a sample onto the sample introduction probe, the sample stage was seeded with a layer of matrix crystallites by depositing 0.5 µl of matrix solution onto it and then, at the first sign of crystal formation (10–15 s after deposition when viewed under a stereo microscope), gently wiping the droplet off with a Kimwipe tissue. Subsequently, a 0.5 µl droplet of analyte/matrix solution (1:4 or 1:9, v/v) was deposited on top of the seed layer and allowed to dry. In some cases, the analyte/matrix deposits were rinsed by gently dipping the sample stage into water before the sample had dried.

Electrospray ionization mass spectrometry

Pneumatically-assisted ESI (ionspray) mass analyses were performed on a Perkin-Elmer Sciex API III Plus triple quadrupole mass spectrometer system. The potential of the electrospray needle was placed at 5000 V or –4100 V to produce positive or negative ions respectively, and the potential of the orifice leading into the mass analyzer was set at 60–110 V or –110 V to extract positive or negative ions, respectively. The flow rates for the nebulizer gas (oxygen) and the curtain gas (nitrogen) were both set at 0.6 l/min for either positive or negative ions. The instrument was mass-calibrated in the positive ionization mode with a mixture of polypropylene glycols.

Samples were injected at 3.3 µl/min with a Harvard syringe pump connected to the electrospray interface via a Rheodyne 8125 injector (5 µl sample injection loop). The synthetic peptide was analyzed in an aqueous solution of 0.1% formic acid/50% methanol; the aminolinker–oligothymidylic acids, the crude reaction mixtures, and the peptide-linker–oligothymidylic acid heteroconjugates were analyzed in a solution of 5 mM ammonium acetate/50% acetonitrile.

The analyzing section of the Sciex API III consists of two quadrupole mass filters separated by an open quadrupole structure that serves as a gas collision chamber. Single stage mass analysis was conducted by scanning the first quadrupole mass filter over the mass-to-charge range of interest (typically *m/z* 400–2200) in ~4 s. Double stage or tandem mass spectrometry was carried out by setting the first quadrupole mass filter to select a doubly charged analyte ion and scanning the second quadrupole mass filter over the mass-to-charge range of interest (*m/z* 50–1300 for peptide, *m/z* 50–2000 for aminolinker-dT_6, and *m/z* 50–2400 for peptide-linker-dT_6) in ~4 s. Fragmentation was induced in the middle quadrupole chamber by introducing argon as a collision gas (mean collision gas density/collision energy: 125/24 eV for peptide; 110/41 eV for aminolinker-dT_6; 185/22 eV for peptide-linker-dT_6).

RESULTS

Synthesis of heteroconjugates

The synthetic peptide **1** was patterned after the sequence of the 11-residue N-terminal segment of Ung-T19 **2**, one of the tryptic peptides of *Escherichia coli* Uracil-DNA glycosylase (Ung) that has been shown to form part of the protein's DNA binding domain (23). The peptide was designed to possess one free carboxyl group by replacing the Asn residue in T19 with an Asp residue. To simplify the synthesis of the peptide–nucleic acid conjugates, the Cys and His amino acids in **2**, whose side chains contain potentially reactive functionalities, were replaced with the less reactive residues Ala and Arg respectively. To prevent side reactions during conjugation of the peptide to the nucleic acid, the N-terminal was blocked by acetylation, and the C-terminal

carboxyl group was amidated. MALDI and ESI mass analyses confirmed that the synthesized peptide had the predicted molecular mass (Table 1), and ESI tandem mass spectrometry confirmed that it had the expected sequence (data not shown).

Ac-Gly-Phe-Phe-Gly-Ala-Asp-Arg-Phe-Val-Leu-Ala-NH_2 (M_r 1240.4) **1**
Gly-Phe-Phe-Gly-Cys-Asn-His-Phe-Val-Leu-Ala **2**

Oligothymidylic acids were used in the synthesis of the heteroconjugates partly because they were available as gifts and partly because they are generally more stable and better behaved in mass spectrometric analyses. The experimentally determined masses of the two oligonucleotide reactants (Table 1) were consistent with the structures NH_2-$(CH_2)_6$-5′-p$(dT)_6$ and NH_2-$(CH_2)_6$-5′- p$(dT)_{10}$ specified by the supplier, Oligos Etc., Inc.

PyBOP was used to catalyze formation of the amide linkage between the carboxyl group on the side chain of the aspartic acid in **1** and the aminolinker–oligothymidylic acid because of its known efficiency in promoting peptide amide bonds during peptide synthesis (28).

Table 1. Average molecular masses of the synthetic peptide, aminolinker–oligothymidylic acids and peptide-linker–oligothymidylic acids

	Molecular mass (Da)		
Compound	Theoretical	MALDI	ESI
Peptide	1240.4	1240.9	1240.4
Aminolinker-dT_6	1942.4	1942.4	1942.0
Aminolinker-dT_{10}	3159.2	3156.6	3158.0
Peptide-linker-dT_6	3164.8	3164.4	3164.3
Peptide-linker-dT_{10}	4381.6	4379.6	4382.0

The theoretical values are calculated from molecular formulas. The experimental values listed under MALDI and ESI are averages over two or more measurements; accuracies (relative to the theoretical values) are within 0.08% for MALDI and 0.02% for ESI.

MALDI mass spectrometry of heteroconjugates

A MALDI mass spectrum of a coupling reaction mixture containing the synthetic peptide and the aminolinker-dT_{10} is shown in Figure 1; a similar spectrum (not shown) was obtained from the reaction mixture containing the aminolinker-dT_6. Formation of the heteroconjugate is indicated in the spectrum by the appearance of a signal corresponding to a compound with a mass equal to the sum of the masses of the peptide and the nucleotide (both of whose signals are present in the spectrum) less the mass of a water molecule, which is lost during the coupling reaction. There is no significant signal in the mass spectrum that might conceivably be construed as a non-covalent complex.

The reactions were also monitored by microbore reverse-phase HPLC. In each case, a conspicuous species eluted between the oligonucleotide and the peptide; these two compounds were collected, lyophilized and mass analyzed by MALDI. Measurement of the masses of these purified compounds was facile and identified them as the two expected peptide–oligothymidylic acids (Table 1). By quantitating absorption in the HPLC eluents at 256 nm and assuming that the extinction coefficient of a heteroconjugate is essentially the same as that of its oligothymidylic acid component, we estimated that the yield of product in both cases was ~5% relative to the amount of aminolinker–oligothymidylic acid used in each of the reactions. Judging from the response of the UV-detector on the microbore system, we estimated that each

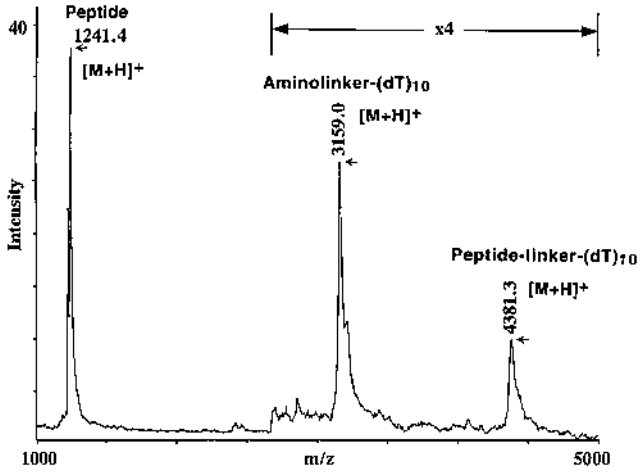

Figure 1. Positive ion MALDI mass spectrum of the reaction mixture used to synthesize peptide-linker-dT_{10}. The signals corresponding to singly protonated forms of the unreacted peptide (*m/z* 1241.4), the unreacted aminolinker-dT_{10} (*m/z* 3159.0), and the hybrid product (*m/z* 4381.3) are labeled in the spectrum.

MALDI mass spectrum was obtained with 10–20 pmol of compound. On our MALDI instrument, which does not operate optimally in the negative ion mode, positive ionization of the nucleopeptide conjugates produced more intense, sharper ion signals than did negative ionization.

ESI mass spectrometry of heteroconjugates

Negative ESI mass analysis of the two crude reaction mixtures (containing primarily synthetic peptide, aminolinker–oligothymidylic acid, PyBOP and peptide-linker–oligothymidylic acid) generated signals corresponding only to the peptide and the aminolinker–oligothymidylic acid (data not shown). Only peptide ions were detected in the positive ESI mode.

Whereas we could not find ion signals corresponding to the heteroconjugates in the ESI mass spectra of the reaction mixtures, we could easily do so in both the negative and positive ESI mass spectra of the HPLC-purified forms. The negative ion spectrum of the purified peptide-linker-dT_6 conjugate is similar in appearance to the positive ion one except sodium adduction is more pronounced in the latter case (Fig. 2). In this instance, the total-ion-current, i.e. the sum of all the ion signals detected in the m/z-range scanned, is more than three times greater in the positive ion mode than the negative. The negative and positive ESI mass spectra of the peptide-linker-dT_{10} are also similar in appearance (Fig. 3). In this case, sodium adduction is quite extensive in the positive ion mode, and the total-ion-current in the positive mode is slightly smaller than in the negative.

ESI tandem mass spectrometry of peptide-linker-dT_6

In the positive ESI mode, tandem mass spectrometry of the doubly charged synthetic peptide (*m/z* 621) indicated fragmentation at every peptide bond in the oligomer (data not shown). In the negative mode, the doubly charged parent of the aminolinker-dT_6 (*m/z* 969.5) fragmented primarily into $[dT\text{-}H_2O\text{-}H]^-$ at *m/z* 303.2 and $[dT]^-$ at *m/z* 321.2 (data not shown). Signals corresponding to PO_3^- (*m/z* 79), [thymine base]$^-$ (*m/z* 125), $[dT\text{-base-}H_2O\text{-}H]^-$ (*m/z* 177.1), and [dT-base-H]$^-$ (*m/z* 195.1) were also observed. Two series of peaks seen in the spectrum, $[dT_n\text{-}H]^-$ and

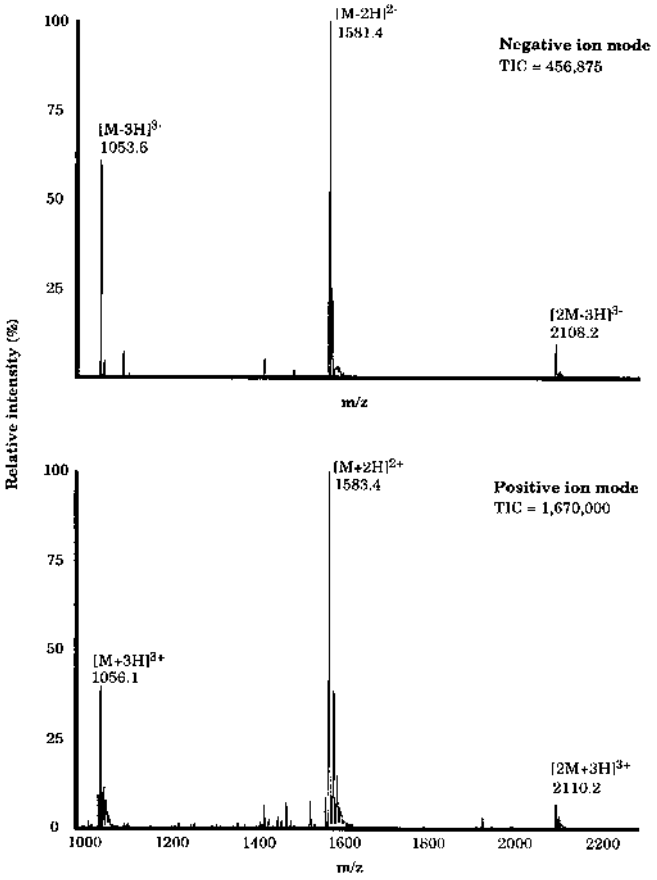

Figure 2. Negative and positive ESI mass spectra of peptide-linker-dT_6 (5 µl of ~50 µM analyte solution; 16 scans per spectrum) following purification by HPLC. The doubly and triply charged ions and the triply charged dimer ion of the complex are labeled in each spectrum. The ion signals adjacent to each mass peak on the high *m/z*-side are due to $[M+n\mathrm{Na}\pm3\mathrm{H}]^{3\pm}$, $[M+n\mathrm{Na}\pm2\mathrm{H}]^{2\pm}$ and $[2M+n\mathrm{Na}\pm3\mathrm{H}]^{3\pm}$ (n = 0–4). The total-ion-current measured in the positive mode was more than three times greater than the total-ion-current measured in the negative mode for the same number of scans.

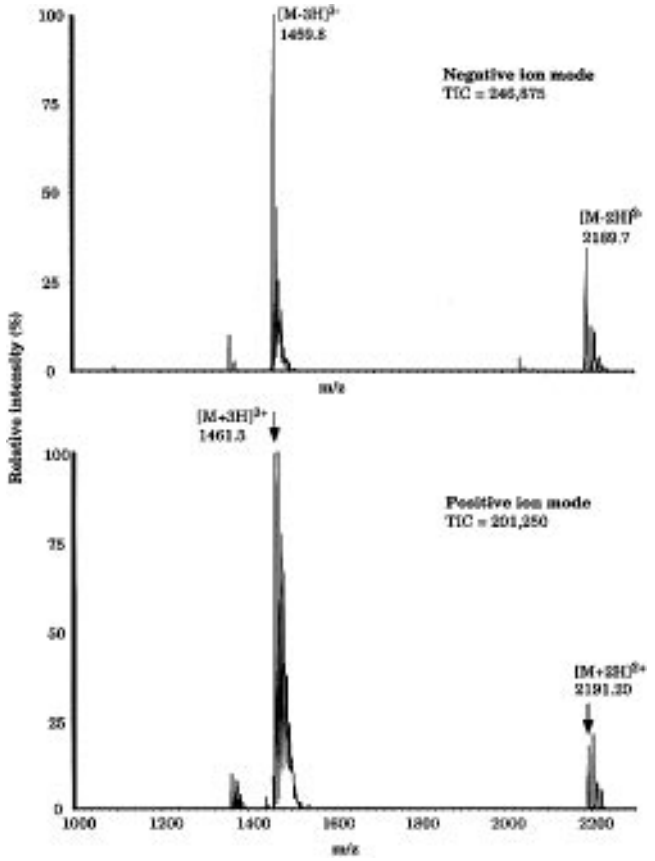

Figure 3. Negative and positive ESI mass spectra of the HPLC eluent corresponding to peptide-linker-dT_{10} (5 µl of ~50 µM analyte solution; 15 scans per spectrum). The doubly and triply charged ions of the complex are labeled in each spectrum. The ion signals adjacent to each mass peak on the high *m/z*-side are due to $[M+n\mathrm{Na}\pm3\mathrm{H}]^{3\pm}$ and $[M+n\mathrm{Na}\pm2\mathrm{H}]^{2\pm}$ (n = 0–6). The total-ion-current measured in the positive mode was only slightly less than the total-ion-current measured in the negative mode for the same number of scans.

$[dT_n\text{-}H_2O\text{-}H]^-$ (n = 2, 3), are specific to the oligothymidylic acid's sequence. Only very weak signals due to larger oligomeric fragments (n = 4 and 5) and no peaks corresponding to ion fragments originating from the aminolinker moiety were observed.

Spectra obtained by tandem mass analysis of the double-negatively charged peptide-linker-dT_6 ion (*m/z* 1581) and the double-positively charged peptide-linker-dT_6 ion (*m/z* 1583) are shown and described in Figures 4 and 5. Using the mass spectra of aminolinker-dT_6 (not shown) as a guide, it is relatively easy to discern several series of negative and positive ion peaks separated by 304 Da, the mass of one thymidylic acid residue, that result from fragmentation of the oligonucleotide. Close inspection of the positive ion spectrum (Fig. 5) in the *m/z*-region below the parent ion's signal reveals peaks corresponding to fragments cleaved from both the N- and C-termini of the peptide.

ESI tandem mass analysis of the peptide-linker-dT_{10} was prevented by the upper mass-to-charge limit (*m/z* 2400) of our instrument.

DISCUSSION

As pointed out in the introduction, the optimum matrix-assisted laser desorption conditions for ionizing the oligopeptide and oligonucleotide components of a covalent hybrid generally clash. In Table 2, six matrices commonly used for MALDI mass analyses of proteins, peptides and oligonucleotides are ranked from top to bottom in the peptide–protein column in order of decreasing suitability. The suitability of these matrices for oligonucleotides is clearly seen to be inverse to that for peptides and proteins. Based on prior work (14,15,23), it seems to us that the ionization of a covalent complex roughly reflects the relative masses of its oligopeptide and oligonucleotide components. Thus on the one hand, sinapinic acid turns out to be the best primary matrix for nucleoprotein heteroconjugates that are 80% or more polypeptide by mass, e.g. phage T4 gene 32 protein UV-crosslinked to dT_{20} (15), *Escherichia coli* transcription termination factor rho photoaffinity labeled with the ATP-analog 4-thio-uridine diphosphate (15), or Ung UV-crosslinked to dT_{20} (23). On the other hand, sample preparation conditions favoring MALDI analysis of the oligonucleotidic component, e.g. 2,4,6-trihydroxyacetophenone in diammonium-hydrogen citrate and acetonitrile (27), are necessary in order to produce the sharpest, most intense ion signals from nucleopeptide conjugates whose peptide component accounts for only 10–40% of the complex's mass, e.g. the nucleopeptidic conjugate from tryptic digests of UV-crosslinked

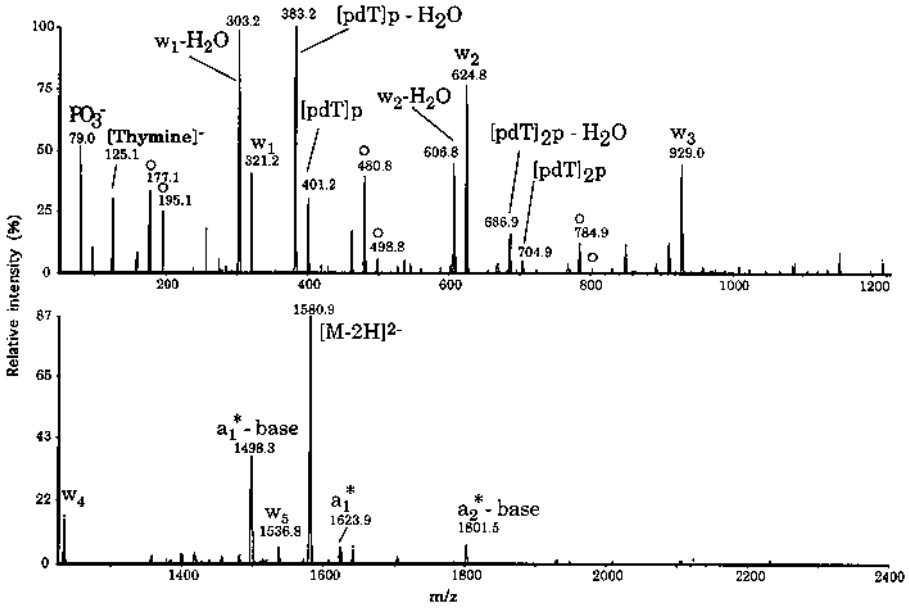

Figure 4. ESI tandem mass analysis of double-negatively charged peptide-linker-dT_6 (*m/z* 1581.4). The main fragment ions from the oligonucleotide part of the heteroconjugate (see Fig. 6 for nomenclature) are PO_3^-, [thymine base]$^-$, w_n (*n* = 1–5), w_n-H_2O (*n* = 1–3), w_n-base (*n* = 1–3), w_n-base-H_2O (*n* = 1–3), $[pdT]_np$ (*n* = 1–2), $[pdT]_np$-H_2O (*n* = 1–2), a_1*, and a_n*-base (*n* = 1–2). The signals corresponding to w_n-base and w_n-base-H_2O (*n* = 1–3) are indicated with open circles (○). No useful peptide-fragment ion signals were observed.

Ung-dT_{20} (23) or the synthetic peptide-linker–oligothymidylic acids used in the present study.

Table 2. Some compounds commonly used as primary matrices for MALDI mass analyses of proteins, peptides and oligonucleotides

Matrix	Peptide and protein	Oligo-nucleotide	Reference
Sinapinic acid	●●●●●	●	(37)
4-hydroxy-α-cyano-cinnamic acid	●●●●●	●	(26)
2,5-dihydroxybenzoic acid	●●●●	●●	(38)
Ferulic acid	●●●	●●●	(37)
2,4,6-trihydroxyacetophenone	●●	●●●●●	(27)
3-Picolinic acid	●	●●●●●	(39)

The general suitability of a given matrix for a given compound class is subjectively indicated on a qualitative scale that ranges from ● (poor) to ●●●●● (excellent).

The interaction between the oligopeptidic and oligonucleotidic parts of a heteroconjugate also seems to manifest itself in the MALDI mass analysis of mixtures. In general, excellent MALDI mass spectra can be produced from mixtures containing large numbers of peptides as is done in peptide mass-mapping of proteins. In our hands, mixtures of peptides and nucleopeptidic conjugates have not exhibited this desirable analytical trait. So far, we have only been able to produce usable mass spectra from mixtures containing two or three heteroconjugates (23), and the strength and definition of the ion signals in the best of these spectra are inferior to those observed in spectra generated from individual, HPLC-purified heteroconjugates. Competition among the various components in the sample for incorporation into the crystalline matrix may be one of the primary factors contributing to the refractory behavior of mixtures of nucleopeptide conjugates.

Conflicting conditions for ionization of oligopeptides and oligonucleotides also exist in electrospray. The heteroconjugates investigated in this study show preference for a minor variation of a solvent system reported for ESI mass analysis of oligonucleotides (29). This seems sensible given that the ratios of peptide-mass to oligonucleotide-mass are ~7:10 for peptide-linker- dT_6 and 4:10 for peptide-linker-dT_{10}. The fact that a 40% decrease in the size of the oligonucleotide component increases the total-ion-current in the positive mode >3-fold over that in the negative mode (compare Fig. 2 with Fig. 3) suggests that the peptide's effect on ionization is disproportionate to its contribution to the size of the heteroconjugate. It should be noted that, for the peptide-linker dT_{10}, most of the signal in the positive ion mode is due to sodium adduction (Fig. 3) and that this effectively reduces sensitivity in this case despite the large total-ion-current. More efficient desalting may remedy this problem.

MALDI mass analysis of a crude reaction mixture proves to be a quick, sensitive and straightforward method for unequivocally confirming formation of a covalent peptide-linker–oligothymidylic acid product (Fig. 1), although it does not provide quantitative information. This capability was used extensively in a study of the DNA binding site of Ung (23), an enzyme that initiates the uracil-excision DNA-repair pathway by removing a uracil-base introduced into DNA (30–32). Results from MALDI mass analyses were combined in this study with concurrent results from Edman sequencing and biological testing to demonstrate that amino acids 58–80 (tryptic peptide 6) and amino acids 185–213 (tryptic peptides 18 and 19) reside at the interface of an Ung–DNA complex. These two domains in the primary sequence of Ung, which are heavily conserved in general, are especially similar to the corresponding amino acid regions in human uracil-DNA glycosylase (UNG 15) (23). Viewing their findings for Ung in light of this fact, Bennett *et al*. ventured that the two

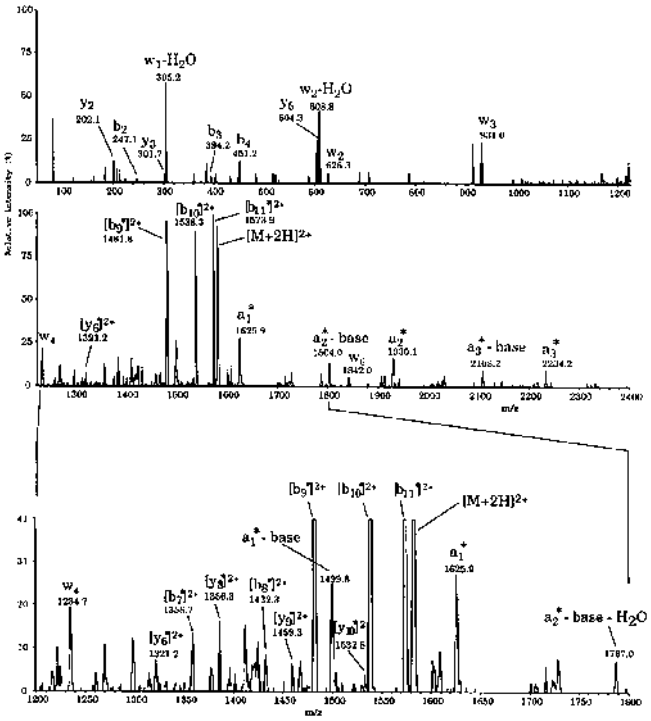

Figure 5. ESI tandem mass analysis of double-positively charged peptide-linker-dT$_6$ ion (*m/z* 1582.9). The inset shows the expanded region *m/z* 1200–1800. Fragment ions from the oligonucleotide part of the heteroconjugate (see Fig. 6 for nomenclature) are w_n (*n* = 2–4), w_n-H_2O (*n* = 1–3), and a_n* (*n* = 1–3). Ions at *m/z* 1499.8, 1804.0 and 2108.2 and at *m/z* 1481.8, 1787.0 and 2091.0 correspond respectively to a_1*-base (*n* = 1–3) and a_1*-base-H_2O (*n* = 1–3) fragments. The labeled fragment ions from the peptide part of the heteroconjugate (see Fig. 6 for nomenclature) are b_n (*n* = 2–4), b_n* (*n* = 7–11), y_n (*n* = 2,3,5) and y_n* (*n* = 6,8–10). Heteroconjugate fragment ions containing the branching point are indicated with an asterisk.

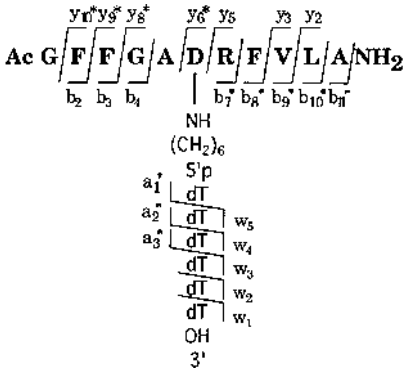

Figure 6. Fragmentation patterns observed in tandem mass analysis of the peptide-linker-dT$_6$. Nucleotide fragment ions are designated as a_n or w_n (40) and peptide fragment ions are designated as b_n or y_n (41,42). The a_n* ions are a_n oligonucleotide-fragment ions with the peptide attached through the C_6-linker; they are unambiguously rationalized in terms of fragmentation from the 5′-end of the oligothymidylic acid, the end bound through the alkyl linker to the peptide. The w_n oligonucleotide-fragment ions can only result from cleavage off the free 3′-end of the oligothymidylic acid. The w_n-H_2O (dT_n) fragments can arise either by cleavage off the 3′-end with subsequent neutral loss of H_2O or by rupture of interior bonds. In the positive ion spectrum (Fig. 5), the three b_n peptide-fragment ions and the four y_n* ions, which are y_n peptide-fragment ions with the linker-dT$_6$ attached, confirm the N-terminal sequence Ac-Gly-Phe-Phe-Gly-Ala- while the three y_n fragment ions and the five b_n* ions, which are b_n peptide-fragment ions with the linker-dT$_6$ attached, confirm the C-terminal sequence -Arg-Phe-Val-Leu-Ala-NH_2. The attachment of the aminolinker-dT$_6$ to the Asp residue is readily deduced from the difference in mass between y_6* and y_5.

UNG 15 amino acid residues Asp145 (corresponding to Asp64 in Ung) and His268 (corresponding to His187 in Ung), whose substitution was recently shown by site-directed mutagenesis studies on the human uracil-DNA glycosylase gene to inactivate UNG 15 (33), reside in the vicinity of the DNA-binding pocket (23). This speculation has since been substantiated by the 3-dimensional structure of UNG 15 obtained from X-ray analysis (34). Another recent X-ray diffraction structure, this one of herpes simplex virus uracil-DNA glycosylase (35), further corroborates the general validity of the Ung-findings by Bennett *et al*. and the utility of the mass spectrometric approach they used to obtain them.

In contrast to the MALDI mass spectra of the reaction mixture containing synthetic heteroconjugates, ESI mass analysis of the crude reaction mixtures in either the positive or negative ionization mode provided no evidence for the formation of the hybrids. Presumably, formation of electrosprayed ions from the small amount of hybrid product (~5%) is suppressed below the limit of detection in the presence of the large amounts of reagents remaining in the reaction mixture.

Molecular masses of purified nucleopeptide heteroconjugates can be readily obtained by either MALDI mass spectrometry or by ESI mass spectrometry. In our hands, ESI requires at least an order of magnitude more sample (50–250 pmol) than does MALDI (5–10 pmol) to produce mass spectra of the linker–oligothymidylic acids and the peptide-linker–oligothymidylic acids, but the ESI mass spectra exhibit better mass resolution and yield better mass accuracy (Table 2) than do their MALDI counterparts.

Mass spectrometry performed in tandem stages is a powerful tool for sequencing biopolymers, especially peptides (36). Under the low-energy collisionally induced dissociation conditions employed in this work, a large number of the double-negatively charged aminolinker-dT$_6$ ions were observed to decompose into dT_n ions (data not shown). Because these fragments could have resulted either from cleavage off the free 3′-end or from rupture of internal bonds, they can not be unambiguously explained in terms of the oligonucleotide's sequence. When tandem mass spectrometry is performed on the double-negatively charged peptide-linker-dT$_6$ (Fig. 4), interpretation of the oligonucleotide's fragmentation is less ambiguous by virtue of the peptide's attachment and, consequently, more is learned about the oligonucleotide's sequence. Tandem mass analysis of the double-positively charged peptide-linker-dT$_6$ also reveals sequence specific information about the oligonucleotide (Figs 5 and 6). The negative ion spectrum of the peptide-linker-dT$_6$ provides no information about the peptide's sequence (Fig. 4), but applied as a background template against the positive ion spectrum, it makes it easier locate the peptide signals in the latter. From the positive ion spectrum (Fig. 5), we can confirm all of the peptide's sequence (Fig. 6). Location of the branching point on the Asp residue is indicated in the spectrum by the low intensity y^*- and y-fragment ion signals (Figs 5 and 6). The intense series of doubly charged b^*-ion peaks in the spectrum (Figs 5 and 6) adds more confidence to the assignment of the branching point because it localizes the covalent bond on the central, three-residue segment of the peptide. These b^*-ions (also present in the fragmentation spectrum of the peptide itself; not shown) probably arise because of the charge carrying capacity of the internal arginyl residue.

Our success with the synthetic peptide-linker dT_6 demonstrates that ESI tandem mass spectrometry can in principle be employed to identify the branching amino acid residue in a nucleopeptide heteroconjugate. It must be noted, however, that results with more heterogeneous nucleopeptides (for example a set of three or four unkowns isolated from a tryptic digest of a protein photochemically crosslinked to some mixed-base oligonucleotide) may vary considerably because fragmentation of a given heteroconjugate in a triple quadrupole system will in general depend on the sequences of its peptide and oligonucleotide. Given our findings, the sequence of a nucleopeptide's oligonucleotide is likely to have greater influence on fragmentation and detection in a tandem mass spectrometer than the sequence of the peptide. Therefore, it may be generally advantageous to reduce the size of a heteroconjugate's oligonucleotide by nucleolytic digestion (chemical or enzymatic) prior to mass analysis.

The tendencies observed with the two synthetic peptide–oligonucleotide conjugates examined in this investigation provide a starting point for future, systematic studies of the relative contributions of oligopeptides and oligonucleotides to the physicochemical properties of heteroconjugates. From a practical perspective, our investigation points to new ways to productively employ mass spectrometry in antisense research and in studies of protein–nucleic acid interactions, namely, monitoring the synthesis or UV-crosslinking of oligopeptide–oligonucleotide hybrids, confirming molecular masses of hybrids, and obtaining residue-specific information pertaining to sites of conjugation.

ACKNOWLEDGEMENTS

This work was supported by grants from the National Institutes of Health Biomedical Research Support (RR07079), from the Medical Research Foundation of Oregon (9146), and from the National Institutes of Health (ES00210 and GM32823). The Sciex API III Tandem Electrospray Mass Spectrometer system was purchased in part through grants from the National Science Foundation (BIR-9214371) and from the Anheuser-Busch Companies. Oligos Etc. of Wilsonville, Oregon generously donated the aminolinker-oligothymidylic acids that made this study possible. O.N.J. was supported by predoctoral fellowships from the Danish Research Academy and the Carlsberg Foundation, Denmark. Dr Seksiri Arttamangkul, College of Pharmacy, provided invaluable advice on the synthesis of the peptide-linker–oligothymidylic acids.

REFERENCES

1 Vartapetian,A.B. and Bogdanov,A.A. (1987) *Prog. Nucleic Acid Res. Mol. Biol.*, **34**, 209–251.
2 Wang,J.C. (1985) *Annu. Rev. Biochem.*, **54**, 665–692.
3 Mathews,C.K. and van Holde,K.E. (1996) *Biochemistry*, 2nd Edn. Benjamin/Cummings, Menlo Park, CA.
4 Smith,K.C. (1962) *Biochem. Biophys. Res. Commun.*, **8**, 157–163.
5 Shetlar,M.D. (1980) In Smith,K.C. (ed.), *Photochemical and Photobiological Reviews*. Plenum Press, New York, Vol. 5, pp. 105–197.
6 Budowsky,E.I. and Abdurashidova,G.G. (1989) *Prog. Nucleic Acid Res. Mol. Biol.*, **37**, 1–65.
7 Williams,K.R. and Konigsberg,W.H. (1991) In Sauer,R.T. (ed.), *Methods in Enzymology—Protein–DNA Interactions*. Academic Press, San Diego, Vol. 208, pp. 516–539.
8 Bayard,B., Bisbal,C. and Lebleu,B. (1986) *Biochemistry*, **25**, 3730–3736.
9 Bashkin,J.K., McBeath,R.J., Modak,A.S., Sample,K.R. and Wise,W.B. (1991) *J. Org. Chem.*, **56**, 3168–3176.
10 Eritja,R., Pons,A., Escarceller,M., Giralt,E. and Albericio,F. (1991) *Tetrahedron*, **47**, 4113–4120.
11 Degols,G., Leonetti,J.–P., Benkirane,M., Devaux,C. and Lebleu,B. (1992) *Antisense Res. Dev.*, **2**, 293–301.
12 Zhu,T., Tung,C.-S., Breslauer,K.J., Dickerhof,W.A. and Stein,S. (1993) *Antisense Res. Dev.*, **3**, 349–356.
13 Bongatz,J.-P., Aubertin,A.-M., Milhaud,P.G. and Lebleu,B. (1994) *Nucleic Acids Res.*, **22**, 4681–4688.
14 Jensen,O.N., Barofsky,D.F., Young,M.C., von Hippel,P.H., Swenson,S. and Seifried,S.E. (1994) In Crabb,J.W. (ed.), *Techniques in Protein Chemistry V*. Academic Press, San Diego, pp. 27–37.
15 Jensen,O.N., Barofsky,D.F., Young,M.C., von Hippel,P.H., Swenson,S. and Seifried,S.E. (1993) *Rapid Commun. Mass Spectrom.*, **7**, 496–501.
16 Karas,M., Bachmann,D., Bahr,U. and Hillenkamp,F. (1987) *Int. J. Mass Spectrom. Ion Processes*, **78**, 53–68.
17 Hillenkamp,F., Karas,M., Beavis,R.C. and Chait,B.T. (1991) *Anal. Chem.*, **63**, 1193A–1202A.
18 Meng,C.K., Mann,M. and Fenn,J.B. (1988) *Zeitschrift für Physik D—Atoms, Molecules and Clusters*, **10**, 361–368.
19 Fenn,J.B., Mann,M., Meng,C.K., Wong,S.F. and Whitehouse,C.M. (1989) *Science*, **246**, 64–71.
20 Wang,R. and Chait,B.T. (1994) *Curr. Opin. Biotech.*, **5**, 77–84.
21 Siuzdak,G. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 11290–11297.
22 Limbach,P.A., Crain,P.F. and McCloskey,J.A. (1995) *Curr. Opin. Biotech.*, **6**, 96–102.
23 Bennett,S.F., Jensen,O.N., Barofsky,D.F. and Mosbaugh,D.W. (1994) *J. Biol. Chem.*, **269**, 21870–21879.
24 Snyder,K.R., Story,S.C., Heidt,M.E., Murray,T.F., DeLander,G.E. and Aldrich,J.V. (1992) *J. Med. Chem.*, **35**, 4330–4333.
25 King,D.S., Fields,C.G. and Fields,G.B. (1990) *Int. J. Pept. Prot. Res.*, **36**, 255–266.
26 Beavis,R.C., Chaudhary,T. and Chait,B.T. (1992) *Org. Mass Spectrom.*, **27**, 156–158.
27 Pieles,U., Zürcher,W., Schär,M. and Moser,H.E. (1993) *Nucleic Acids Res.*, **21**, 3191–3196.
28 Coste,J., Le–Nguyen,D. and Castro,B. (1990) *Tetrahedron Lett.*, **31**, 205–208.
29 Stults,J.T. and Marsters,J.C. (1991) *Rapid Commun. Mass Spectrom.*, **5**, 259–263.
30 Lindahl,T., Ljungquist,S., Siegert,S., Nyberg,B. and Sperens,B. (1977) *J. Biol. Chem.*, **252**, 3286–3294.
31 Dianov,G., Price,A. and Lindahl,T. (1992) *Mol. Cell. Biol.*, **12**, 1605–1612.
32 Mosbaugh,D.W. and Bennett,S.W. (1994) *Prog. Nucleic Acid Res. Mol. Biol.*, **48**, 315–370.
33 Krokan,H.E., Slupphaug,G., Alseth,I., Kavli,B., Haug,T., Olsen,L.C. and Helland,D.E. (1993) New York Academy of Sciences, Burlington, VT, (abstr.).
34 Mol,C.D., Arvai,A.S., Slupphaug,G., Kavli,B., Alseth,I. and Krokan,H. (1995) *Cell*, **80**, 869–878.
35 Savva,R., McAuley-Hecht,K., Brown,T. and Pearl,L. (1995) *Nature*, **373**, 487–493.
36 Biemann,K. (1992) *Annu. Rev. Biochem.*, **61**, 977–1010.
37 Beavis,R.C. and Chait,B.T. (1989) *Rapid Commun. Mass Spectrom.*, **3**, 432–435.
38 Strupat,K., Karas,M. and Hillenkamp,F. (1991) *Int. J. Mass Spectrom. Ion Processes*, **111**, 89–102.
39 Wu,K. J., Steding,A. and Becker,C.H. (1993) *Rapid Commun. Mass Spectrom.*, **7**, 142–146.
40 McLuckey,S.A., Van Berkel,G. J. and Glish,G.L. (1992) *J. Am. Soc. Mass Spectrom.***3**, 60–70.
41 Roepstorff,P. and Fohlman,J. (1984) *Biomed. Mass Spectrom.*, **11**, 601.
42 Biemann,K. (1990) In McCloskey,J.A. (ed.), *Methods in Enzymology*. Academic Press, Inc., San Diego, Vol. 193, pp. 886–887.