Published in final edited form as:

J Chromatogr B Analyt Technol Biomed Life Sci. 2007 December 15; 860(2): 145–152.

Comparison of ESI MS Interfaces for the Analysis of UV-crosslinked Peptide-Nucleic Acid Complexes

Philip R. Gafken^{1,*}, Catalin E. Doneanu^{2,#,*}, Samuel E. Bennett³, and Douglas F. Barofsky⁴

¹ Proteomics Facility, Fred Hutchinson Cancer Research Center, Seattle, WA 98109

² Department of Chemistry, Oregon State University, Corvallis, OR 97331

³ Department of Environmental & Molecular Toxicology, Oregon State University, Corvallis, OR 97331

⁴ Environmental Health Sciences Center, Oregon State University, Corvallis, OR 97331

Abstract

In this report, the effectiveness of high performance liquid chromatography (HPLC) in conjunction with electrospray ionization mass spectrometry (ESI-MS) is examined as a tool for identifying the sites of crosslinking in a protein that has been photoreacted with a non-photolabeled oligonucleotide. ESI-MS and MALDI-MS analyses preceded by off-line microflow and nanoflow HPLC, on-line microflow HPLC/ESI, and on-line nanoflow HPLC/ESI interfaces were performed in order to determine their relative effectiveness in separating mixtures of nucleopeptides and identifying sites of crosslinking on the individual components. The characteristics of these four techniques as well as possibilities for improving the analysis of nucleopeptides by ESI MS are compared and discussed.

INTRODUCTION

The use of photochemical crosslinking as a tool for locating the sites of interaction between proteins and nucleic acids dates back to the early 1960s [1]. The UV light fixes the interaction between a protein and a nucleotide by inducing a “zero-length”, covalent bond between the two biopolymers [2]. Although photochemical crosslinking for the formation of the covalent bonds is not fully understood, the technique has been exploited to investigate both static and dynamic protein-nucleic acid complexes [3,4]. In the 1990s, Barofsky and coworkers demonstrated that matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (MS) could be effectively employed to identify the DNA-binding domains in proteins photochemically crosslinked to oligonucleotides [5,6]. This strategy of combining photochemical crosslinking and mass spectrometry has been used by several investigators to characterize the interface between protein and oligonucleotide in a variety of heteroconjugates. Recently, this body of work, as well as the various strategies that have emerged to photochemically produce and mass spectrometrically analyze protein/peptide-nucleic acid heteroconjugates, was comprehensively reviewed by Steen and Jensen [7].

Address reprint requests to Dr. Douglas F. Barofsky, Oregon State University, Department of Chemistry, 153 Gilbert Hall, Corvallis, OR, 97331-4003, Phone: 541 737-1769, Fax: 541 737-8117, e-mail: doug.barofsky@oregonstate.edu.

#Current address: Waters Corporation, 34 Maple Street, Milford, MA 01757

* Authors contributed equally to this work

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

The potential for sequencing peptide-nucleic acid complexes via electrospray ionization (ESI) tandem mass spectrometry (MS/MS) was first demonstrated using a synthetic peptide chemically linked to the oligonucleotide dT₆ [8]. Results from this early study showed that ESI MS/MS could sequence the peptide portion of the complex, sequence the oligonucleotide portion of the complex, and identify the chemically crosslinked amino acid. Apparently, the first application of ESI MS/MS to the analysis of a peptide-oligonucleotide heteroconjugate derived from a photochemically crosslinked protein-DNA complex was made in 1999 [9]. Subsequently, other laboratories have used it to characterize the interface between peptide and oligonucleotide in heteroconjugates formed either by photochemical crosslinking [10–12] or by chemical crosslinking [13,14]. Interestingly, in those mass spectrometric studies of photochemical crosslinking where the nucleic acid substrate was photolabeled (usually with either 5-bromo- or 5-iodouracil) [12,15], crosslinking to the protein was found to occur at only a single amino acid residue regardless of the method of mass analysis. In those studies where the nucleic acid substrate was not photolabeled, crosslinking to the protein was found to occur at a number of amino acid sites [10,11].

In the present study, complex mixtures of dT_n (n=2–6) nucleopeptides were isolated and characterized by microflow/nanoflow chromatography coupled off-line or on-line with ESI and MALDI mass spectrometry. The goal of this survey was to find which chromatographic configuration allowed for a complete mass spectrometric characterization of the nucleopeptide mixture.

EXPERIMENTAL SECTION

Chemicals

Acetic acid (AA), trifluoroacetic acid (TFA), formic acid (FA), diammonium hydrogen citrate, Tris-HCl, and 2,4,6-trihydroxyacetophenone (THAP), were purchased from Sigma Chemical Co. (St Louis, MO). HPLC grade acetonitrile, ammonium acetate, ammonium bicarbonate, glycerol, dithiothreitol, sodium chloride, EDTA, isopropanol and methanol were supplied by Fisher Scientific (Pittsburgh, PA). Water was generated with a Milli-Q water purification system (Millipore Corp., Bedford, MA). Oligodeoxythymidylate dT₂₀, was synthesized by the Biopolymer Core Facility, at the University of Maryland at Baltimore and further purified by native polyacrylamide (15%) gel electrophoresis. Ung was overexpressed and purified as described by Bennett et al. [16] with modifications described by Sanderson and Mosbaugh [17].

UV Crosslinking

Mixtures of crosslinked nucleopeptides were produced by placing 36 nmol of purified Ung and a 3-fold molar excess of dT₂₀ in a quartz cuvette (4 mL) and adding DAB buffer (30 mM Tris-HCl, 50 mM NaCl, 1 mM dithiothreitol, 5 % (w/v) glycerol, pH=7.4) to bring the final volume to 1 mL; placing the cuvette on ice for 15 minutes, and finally, laying the cuvette lengthwise on a bed of ice in a Stratalink 1800 (Stratagene Cloning Systems, La Jolla, CA) and irradiating ($\lambda_{\text{max}} = 254 \text{ nm}$) the reaction mixture for 15 min. The irradiated solution was either used immediately or stored at -80°C .

Isolation of nucleopeptide mixtures

The Ung-oligonucleotide complexes produced by photocrosslinking were isolated using denaturing polyacrylamide gel electrophoresis (PAGE) and digested in the gel with trypsin (36 ng/ μL) as described by Shevchenko et al. [18]. The peptide $\times$ dT₂₀ fragments (nucleopeptides) resulting from this procedure were extracted and purified as described in detail by Gafken [19]. Briefly, the aqueous digestion solution was removed and saved. The tryptic peptides remaining in the gel were extracted in two stages respectively into FAPH (50% formic acid,

25% acetonitrile, 15% isopropyl alcohol, and 10% water) and acetonitrile. The FAPH and acetonitrile were combined, concentrated to dryness by vacuum centrifugation, and re-suspended with 250 μ L of 10 mM Tris-HCl (pH 7.2), 1 mM EDTA, and 100 mM sodium chloride (TE-100 buffer). The aqueous digestion solution saved earlier was diluted in TE-100 buffer and mixed with the re-suspended organic extract. This composite solution of tryptic proteolysis products was loaded onto a NAC-52 anion exchange cartridge (Life Technologies Inc., Grand Island, NY); the cartridge was washed with 5 mL of TE-100; and the nucleopeptides were eluted with 1 mL of TE-1000 (10 mM Tris-HCl (pH 7.2), 1 mM EDTA, and 1000 mM NaCl) directly into a Centricon-3 (Millipore, Billerica, MA) cartridge for centrifugal desalting. The washed sample was concentrated to 200 μ L and then transferred to a 750- μ L Eppendorff tube and further concentrated to about 10 μ L by vacuum centrifugation. Nuclease P1 (Amersham Biosciences, Piscataway, NJ) was diluted with 50 mM ammonium acetate to 0.04 activity units/ μ L. The isolated peptide $\times$ dT₂₀ complexes were digested by mixing 1 μ L of nuclease P1 solution with 1 μ L of nucleopeptide sample and allowing the reaction to proceed at 37 °C for 4 h.

Off-line microflow HPLC

Chromatography was performed off-line on a system made up of an ABI 140B syringe pump (Applied Biosystems, Foster City, CA), a flow splitter (Upchurch Scientific, Murrieta, CA), a 0.32 mm (i.d.) $\times$ 250 mm column packed in-house with 5- μ m, 150 Å pore, BetaBasic C₁₈ packing material (Keystone Scientific/Thermo Electron Corporation, San Jose, CA), and an ABI 759 UV detector (Applied Biosystems, Foster City, CA) equipped with a 35-nL dead volume, 8-mm path-length flowcell (LC Packings, Sunnyvale, CA). All connections in the system were made with 50 μ m (i.d.) silica tubing (Polymicro Technologies, Phoenix, AZ). The solvents used in the mobile phase were 5mM ammonium acetate (A) and 100% acetonitrile (B); solvent was delivered at 5 μ L/min while linearly increasing the concentration of B from 5% to 50% over an interval of 45 min. Fractions containing species detected at 217 nm were collected off-line for subsequent identification by MALDI and nanospray MS.

Matrix-assisted laser-desorption ionization MS

MALDI MS was performed on a custom-built, delayed extraction, time-of-flight instrument [5]. The matrix solution containing 50% acetonitrile and 50 mM diammonium hydrogen citrate was saturated with 2',4',6' trihydroxyacetophenone (THAP). Nucleopeptide sample and matrix were mixed in a ratio of 1:2 (v:v), and 1 μ L of this mixture was applied to a pre-crystallized layer of THAP on a stainless steel probe. After allowing the spot to dry, it was rinsed with deionized water and again allowed to dry. The ion source potential was set at +20 kV, and the extraction delay was set at 500 ns. Mass spectra were produced using *MoverZ* software (Genomic Solutions Inc, <http://www.genomicsolutions.com>) to sum the signals generated from 30 individual laser pulses.

Nanoelectrospray (nano ESI) MS

The nanospray analyses were performed on an API III triple quadrupole mass spectrometer (MDS Sciex, Concord, Ontario) in which the commercial electrospray source was replaced with a nanospray source constructed by the Protein and Peptide Group in the European Molecular Biology Laboratory, (Heidelberg, Germany) [20]. Nanospray needles were purchased from Protana Inc. (Toronto, Ontario, Canada). The flow rate of the nitrogen curtain gas was 0.6 μ L/min; the needle voltage was 700 V; the interface plate voltage was 100 V; and the orifice potential was 80 V. Each sample collected off the HPLC system was adjusted to 1% formic acid and 40% acetonitrile; 0.6 μ L of adjusted sample was loaded into a nanospray needle. MS spectra were acquired by stepping the first quadrupole in m/z -increments of 0.3 over 3 s scan-intervals.

On-line microflow HPLC/ESI MS

The chromatography system comprised an ABI 140B syringe pump (Applied Biosystems, Foster City, CA), a flow splitter (Upchurch Scientific, Murrieta, CA), and a 0.17 mm (i.d.) $\times$ 100 mm column packed in-house with 5- μ m, 100-Å pore, Luna C₁₈ packing material (Phenomenex, Torrance, CA). The solvents used in the mobile phase were 0.1% acetic acid, 0.01% trifluoroacetic acid (A) and 0.1% acetic acid, 0.01% trifluoroacetic acid in acetonitrile (B); solvent was delivered at 2 μ L/min while linearly increasing the concentration of B from 5% to 50% B over an interval of 30 min. The HPLC column was connected by 30 μ m (i.d.) silica tubing to the custom built ESI source of a Finnigan LCQ ion trap mass spectrometer (Thermo Electron Corporation, San Jose, CA). The ESI sprayer was a 2.5 cm long, 35 gauge, stainless steel hypodermic needle mounted on an XY-manipulator; the tip of the needle was positioned approximately 5 mm from the mass spectrometer's inlet capillary. During operation, the spray-needle was held at 2.3 kV, the temperature of the capillary inlet was maintained at 180 °C, the capillary potential at 46 V, and the tube-lens offset-potential at 30 V. Mass spectra were acquired between m/z 400 and 2000 with a maximum injection time set at 50 ms.

Off-line nano-LC/ESI MS

A nano-LC column was prepared by packing a 15 cm length of 360 μ m o.d. $\times$ 75 μ m i.d. fused-silica (Polymicro Technology, Phoenix, AZ) with 5 μ m, 300 Å pore, Luna C₁₈ silica gel particles (Phenomenex, Torrance, CA) using a pressurized bomb method described by Kennedy and Jorgensen [21]. In short, the end of the silica capillary that is to become the inlet was slipped through a seal into the bomb containing a slurry of isopropanol and the packing material (25 mg/mL); the outlet of the column was inserted into a Valco (Houston, TX) microbore end-fitting containing a 2 μ m Valco metallic screen that acted as a temporary outlet frit. The packing material was forced into the capillary by pressurizing the bomb to 1,500 psi; about 2 h were required to pack a 15 cm long column. After the gas pressure was slowly released from the bomb, the inlet of the column was connected to an HPLC pump and flushed with acetonitrile for 2 h and Milli-Q water for another 2 h. A fused-silica splicer (Fujikura, Japan) was used to prepare the final outlet frit while the column was still wet. An exponential dilution method described in detail elsewhere [22] was used to produce gradient separations. Briefly, a ZDV PEEK union (Upchurch Scientific, Murrieta, CA) drilled out to an internal volume of 3 μ L and the 10 μ L inlet port of the micro-injection valve (Rheodyne Model 8125, Rohnert Park, CA) were used as the mixing chambers. Two Kratos Spectroflow 400 HPLC pumps (Kratos Analytical, Chestnut Ridge, NY) were used to independently deliver solvent A (0.1 % AA plus 0.01 % TFA) and solvent B (0.1 % AA plus 0.01 % TFA in 50% acetonitrile) to the mixing chambers. PEEK tubing (Upchurch Scientific, Murrieta, CA) with 60 μ m i.d. was used to connect the pumps to the filter, splitter, mixing chamber, and injector. The inlet of nanocolumn was attached directly into the injector's outlet port. The outlet of the nanocolumn was connected via a 30 μ m i.d. $\times$ 20 cm fused-silica transfer line to a capillary electrophoresis (CE) unit (Beckmann P/ACE System 2210, Fullerton, CA) which provided UV detection and data acquisition. The nanocolumn was attached to the transfer capillary with a short piece of Teflon tubing; the length of the transfer capillary from the outlet of the nanocolumn up to the detector was 13 cm. A small diameter hole was drilled in the upper part the CE unit's cartridge (the housing that holds the capillary in place) to allow the packed capillary column to be connected to the transfer line inside the cartridge. A small hole, aligned with the hole in the cartridge, was also drilled in the upper right part of the cartridge interface of the CE unit.

Sample was displaced from the injection valve's loop (5 μ L) onto the nanocolumn by forcing solvent A through the system at 0.8 μ L/min (inlet pressure = 200 bar) for ~15 min, and it was separated on the column with an exponential gradient produced by forcing solvent B through the system at 0.3 μ L/min (inlet pressure = 120 bar). Fractions were collected manually for subsequent analyses by MALDI and nanoelectrospray MS. This was performed by adding 1

μL of solvent A to a vial and collecting 2 μL of eluate over a 2-minute period to bring the final volume in the vial to 3 μL .

On line nano-LC/ESI MS

Analyses were performed on the LCQ ion trap mass spectrometer operated with the same experimental parameters described in a previous subsection. A 360 μm (o.d.) $\times$ 75 μm (i.d.) $\times$ 40 cm fused-silica capillary with a 15 μm tip and an integral frit (New Objective "PicoFrit", Cambridge, MA) was packed with 5 μm , 300 Å pore, Luna C₁₈ particles. The column was connected directly to the 8125 Rheodyne injector placed in a plastic box, and the spraying voltage (2.5 kV) was applied to the injector after loading the sample. Elution gradients were produced using the exponential dilution setup described in the preceding subsection by mixing solvent A (0.1 % AA plus 0.01% TFA) and solvent B (0.1 % AA plus 0.01% TFA in 40 % acetonitrile).

RESULTS

Off-line microflow HPLC/MALDI MS

Prior to analysis by ESI MS, an aliquot of the nuclease P1 digest of the tryptic nucleopeptides was surveyed by MALDI MS to gauge the separation efficiency. The resulting mass spectra revealed that the sample contained a number of complexes of the form peptide $\times$ dT_n (n = 2–6), i.e. oligodeoxyribothymidylic acids of various lengths crosslinked to certain tryptic peptides (data not shown). Separation of the peptide-nucleic acid mixture by microflow-HPLC produced two late-eluting fractions (Figure 1A). The MALDI mass spectrum of Fraction I (Figure 1B) shows two series of singly charged ion species. The members of the dominant series (m/z 1796.0, 2100.2, 2404.4, 2708.6, and 3012.8) correspond respectively to the polydeoxyribothymidylic acids dT₂ through dT₆ crosslinked with the Ung-peptide T18 [184-APHPSPLSAHR-194]. The signals in the weaker series (m/z 2020.1, 2324.3, 2628.4, and 2932.7) correspond respectively to dT₃ minus a phosphate through dT₆ minus a phosphate crosslinked with T18. A signal due to the free peptide T18 itself also appears in the spectrum at m/z 1169.6

The MALDI mass spectrum of Fraction II (Figure 1C) shows nucleopeptide ion signals at m/z 2315.3, 2485.5, and 2703.8 corresponding to dT₂ crosslinked respectively with T6₍₋₇₎, T6₍₋₅₎, and T6₍₋₃₎ (where these three peptides are hydrolysis products of T6 [57-VVILGQDPYHGPGQAHGLAFSVR-79] from which seven, five, and three C-terminal amino acids, respectively, have been lost during sample preparation), at m/z 3312.2 corresponding to dT₄ crosslinked with T6₍₋₃₎, and at m/z 2570.5, 2874.7, and 3178.9 corresponding respectively to dT₂, dT₃, and dT₄ crosslinked with T11 [129-AGQAHSHASLGWETFTDK-146].

Off-line microflow HPLC/nanoESI MS

Nanospray analyses of the Fractions I and II (Figure 2) produced results similar to those obtained by MALDI MS. The Q1 mass spectrum of Fraction I (Figure 2A) exhibits mass peaks at m/z 700.7, 802.1, 903.5, and 1004.9 corresponding to triply charged T18 $\times$ dT_n with n equal to 3 through 6 respectively. Another series of triply charged T18 $\times$ dT_n ions (n = 4–6) in which the nucleotide has lost a phosphate group is seen in the spectrum at m/z 775.4, 876.8, and 978.2 respectively. The nanospray Q1 mass spectrum of Fraction II (Figure 2B) contains mass peaks corresponding to T6₍₋₇₎ $\times$ dT₂, T6₍₋₅₎ $\times$ dT₂, T11 $\times$ dT₃, and T11 $\times$ dT₄ (m/z 772.4, 829.2, 958.9, and 1060.3 respectively). The mass spectra from both HPLC fractions indicate that significant cation-adduction occurred in the nanospray process. By contrast, no evidence of cation-adduction was found in the MALDI mass spectra.

Off-line nanoflow HPLC/MALDI MS

The nano-LC/UV chromatogram of the nucleopeptide sample digested with nuclease P1 (Figure 3A) indicates the presence of two relatively low abundant species with retention times of roughly 13 and 17 minutes. The baseline absorbance in this chromatogram gradually decreases with increasing time as acetonitrile from solvent B (50 % acetonitrile with 0.1 % TFA) replaces water from solvent A (0.1 % TFA in water); the sensitivity is low (compared with a conventional UV-detector) because of the nanoflow HPLC system's short path length (30 μ m). The MALDI spectrum of Fraction 1 (Figure 3B) shows peaks corresponding to T18 $\times$ dT₂ and T18 $\times$ dT₃ (m/z 1796.1 and 2100.3 respectively), and the MALDI spectrum of Fraction 2 (Figure 3C) contains an intense peak corresponding to T6₍₋₅₎ $\times$ dT₂ (m/z 2485.5).

On-line microflow HPLC/ESI MS

The total ion chromatogram of the nucleopeptide mixture produced by on-line microflow HPLC/ESI (Figure 4A) exhibits peaks corresponding to a large number of compounds. Despite the limited degree of separation evident in the chromatogram, nucleopeptides were found in the eluate coming off the column at 21.5, 23.6, and 24.9 minutes. The mass spectra recorded at these retention times (Figures 4B-D) contain prominent peaks corresponding to doubly and triply protonated T18 $\times$ dT₂ (m/z 898.5 and 599.3 respectively), T11 $\times$ dT₂ (m/z 1285.7 and 857.5 respectively), and T6₍₋₃₎ $\times$ dT₂ (m/z 1352.4 and 901.9 respectively).

On-line nanoflow HPLC/ESI MS

The reconstructed nanoflow LC/ESI-MS chromatogram (Figure 5) and the corresponding mass spectra (not shown) indicate that the nucleopeptides T6₍₋₃₎ $\times$ dT₂, T6₍₋₅₎ $\times$ dT₂, T6₍₋₇₎ $\times$ dT₂, T11 $\times$ dT₂, and T18 $\times$ dT₂ were clearly separated from a nuclease P1 digested nucleopeptide sample. Each of the five nucleopeptide peaks observed in the chromatogram were generated by plotting only the doubly charged molecular ion signals for the species mentioned above.

DISCUSSION

The ultimate outcome of a mass spectrometric analysis of a peptide-nucleic acid complex is identification of those amino acids that are crosslinked to nucleic acids. Such data provides insight into the mechanism by which the protein interacts with its nucleic acid substrate. The work presented here was part of a biochemical project that was conducted to map the DNA-binding domain of *E. coli* uracil-DNA glycosylase.

Uracil-DNA glycosylase is a monomeric, 228 amino acid polypeptide ($M_r = 25,562$ Da) which initiates the uracil-excision DNA repair pathway *in vivo* by cleaving the N-glycosylic bond between uracil and deoxyribose in DNA [23]. Ung shows a two-fold substrate specificity for uracil residues in single-stranded as opposed to double-stranded DNA [23]. While all amino acids have comparable UV cross-linking potentials, nucleic acids have distinctly different cross-linking potentials; specifically, thymidine and uridine respectively are by far the most efficient deoxynucleotide and ribonucleotide cross-linkers [3]. The UV-crosslinking specificity of polydeoxythymidine towards amino acids is so low that virtually any of a DNA-binding protein's amino acids in contact with a thymine base can take part in crosslinking. This lack of specificity increases the number of crosslinks that can form within the protein-nucleic acid interface.

Previously when oligodeoxythymidine dT₂₀ was used as a DNA substrate, three tryptic peptides (T6, T11 and T18) were shown to be involved in UV crosslinking [6,10,19]. This finding relied on the isolation of crosslinked peptide (nucleopeptides) by anion exchange chromatography following the in-gel trypsin digestion of the nucleoprotein complex. Altogether, these previous studies revealed five dT₂₀ nucleopeptides by MALDI mass

spectrometry in a fraction eluted with 1 M TE buffer from the anion exchange column (spectrum not shown). Two of the nucleopeptides were identified as T11xdT₂₀ and T18xdT₂₀, while the other three contained the T6 peptide minus 3, 5 and 7 amino acids from the Ung C-terminal. Nuclease P1 digestion, employed to remove most of the oligonucleotide portion of these complexes and thereby enhance the nucleopeptide signal in positive ESI or MALDI, compounded sample complexity. In the present study, complex mixtures of dT_n (n=2–6) nucleopeptides were isolated and various separation configurations coupled either off-line or on-line with ESI and MALDI mass spectrometry were surveyed. The goal of this survey was to find which configuration allowed for a complete mass spectrometric characterization of a nucleopeptide mixture. The small sample quantities isolated from the crosslinking experiments required the investigation of high sensitivity ESI configurations, namely nanospray/, micro-LC/, and nano-LC/ESI MS. These offline and on-line microflow/nanoflow LC/MS interfaces were evaluated for their ability to provide sample separation and mass spectrometric detection of the five dT₂ nucleopeptides known to be present.

Initial fractionation of the dT_n nucleopeptide mixture was performed by off-line microflow chromatography and two fractions (labeled I and II in the LC/UV chromatogram shown in Figure 1A) were subsequently analyzed by MALDI and nanoelectrospray MS. The MALDI spectrum of the first fraction revealed the presence of five T18 nucleopeptides with the general formula T18xdT_n (n=2–6), as well as four T18 nucleopeptides (n=3–6) without a phosphate group. Fraction II, isolated during the same preparative run, contained seven nucleopeptides, four of which derived from the T6 peptide and three from T11. Clearly, the five dT₂ nucleopeptides in question could not be isolated by off-line microflow LC fractionation.

The reported advantages of nanoelectrospray tandem mass spectrometry, in particular the technique's subpicomole sensitivity for peptides and facile acceptance of submicroliter volumes of sample [25,26], made it an attractive candidate for analyzing isolated nucleopeptides. However, we found that HPLC enriched fractions of nucleopeptides produced very complicated nanospray spectra (Figure 2A-B). The complexity of these spectra arises partly from the multiple charge states of the individual nucleopeptides that are characteristic of electrospray ionization. In general, the charge states of a nucleopeptide were between +2 and +4, usually with the +3 charge state being most abundant. Additional complexity in these spectra came from adduction primarily of sodium and potassium. Leaching from the nanospray needles by the acidic HPLC fractions could explain the abundant presence of the latter in the nanospray spectra. The combination of multiply charged states and cation adduction greatly reduced the technique's sensitivity for single nucleopeptide ion species.

The utility of peptide fractionation by nanoscale LC before MALDI or nano-ESI has been reported [22]. Even though the sensitivity of the UV detection used for guiding the collection of fractions containing nucleopeptides was limited in the present study by a short path-length (30 μ m), this mode of sample fractionation allowed relatively pure dT₂ nucleopeptide species, viz. T18 x dT₂ and T6₍₋₅₎x dT₂ (Figures 3B and 3C) to be isolated. However, the T11xdT₂ nucleopeptide was not detected using this approach.

The on-line LC/ESI-MS approaches employed in this study fully revealed the complexity of the isolated dT_n nucleopeptides mixtures. A typical micro-LC/ESI base-peak chromatogram is presented in Figure 4A. Even though the majority of the peaks present in this chromatogram do not correspond to predicted dT₂ nucleopeptides, three of them do (Figure 3B through C). The charge state distribution of the nucleopeptide species observed during micro-LC/ESI/MS experiments are similar to those observed for nanospray ESI experiments (charge states distributed from +2 to +4, with the +3 charge state being most abundant), but cation adduction is not observed. All three of the dT₂ nucleopeptides (T6, T11 and T18) that define the DNA binding domain of Ung were completely separated using this online approach.

Nanoscale LC/MS analysis improved sensitivity and chromatographic resolution sufficiently to provide a comprehensive nucleopeptide analysis of all dT₂ species. ESI-MS spectra of the five dT₂ nucleopeptides known to exist in the sample were acquired during a 40 min chromatographic run performed on a 40 cm long nanoscale column (Figure 5) with each nucleopeptide producing a distinct chromatographic peak 15–20 seconds wide. Furthermore, this method facilitated the acquisition of collision-induced dissociation spectra (CID) for all of the dT₂ nucleopeptides, thus allowing the individual amino acids responsible for the interaction between Ung and the DNA substrate to be located in a single chromatographic run.

In summary, Table I shows the results obtained from each LC/MS method, indicating which nucleopeptides were detected by each approach.

CONCLUSIONS

In general, on-line LC mass spectrometric analysis of nucleopeptides is considerably less tedious and time consuming than off-line approaches and minimizes cation adduction associated with nanoESI MS. In this study, the nanoscale LC format of LC/MS-ESI proved to be the best method for characterizing nucleopeptides, by virtue of being able to analyze an entire mixture in a single analysis.

References

1. Alexander P, Moroson H. Nature 1962;194:882. [PubMed: 13860354]
2. Shetlar, MD. Photochemical and Photobiological Reviews. Smith, KC., editor. Plenum Press; New York: 1980. p. 105
3. Hockensmith, JW.; Kubasek, WL.; Vorachek, WR.; Evertsz, EM.; von Hippel, PH. Protein-DNA Interactions. Saurer, RT., editor. Academic Press; San Diego: 1991. p. 211
4. Williams, KR.; Konigsberg, WH. Protein-DNA Interactions. Saurer, RT., editor. Academic Press; San Diego: 1991. p. 516
5. Jensen ON, Barofsky DF, Young MC, von Hippel PH, Swenson S, Seifried SE. Rapid Comm Mass Spectrom 1993;7:496.
6. Bennett SE, Jensen ON, Barofsky DF, Mosbaugh DW. J Biol Chem 1994;269:21870. [PubMed: 8063831]
7. Steen H, Jensen ON. Mass Spectrom Rev 2002;21:163. [PubMed: 12476441]
8. Jensen ON, Kulkarni S, Aldrich JV, Barofsky DF. Nucleic Acids Res 1996;24:3866. [PubMed: 8871569]
9. Golden MC, Resing KA, Collins BD, Willis MC, Koch TH. Protein Sci 1999;8:2806. [PubMed: 10631998]
10. Gafken, PR.; Mosbaugh, DW.; Barofsky, DF. Proceedings of the 48th ASMS Conference on Mass Spectrometry and Allied Topics; Long Beach, CA. 2000. p. 55
11. Doneanu CE, Gafken PR, Bennett SE, Barofsky DF. Anal Chem 2004;76:5667. [PubMed: 15456284]
12. Steen H, Petersen J, Mann M, Jensen ON. Protein Sci 2001;10:1989. [PubMed: 11567090]
13. Deterding LJ, Prasad CR, Mullen GP, Wilson SH, Tomer KB. J Biol Chem 2000;275:10463. [PubMed: 10744736]
14. Rieger RA, McTigue MM, Kycia JH, Gerchman SE, Grollman AP, Iden CR. J Am Soc Mass Spectrom 2000;11:505. [PubMed: 10833024]
15. Qin J, Chait BT. Anal Chem 1997;69:4002. [PubMed: 9322437]
16. Bennett SE, Schimerlik MI, Mosbaugh DW. J Biol Chem 1993;268:26879. [PubMed: 8262921]
17. Sanderson RJ, Mosbaugh DW. J Biol Chem 1996;271:29170. [PubMed: 8910574]
18. Shevchenko A, Wilm M, Vorm O, Mann M. Anal Chem 1996;68:850. [PubMed: 8779443]
19. Gafken, PR. PhD Thesis. Oregon State University; 2000. Characterization of UV-crosslinked protein-nucleic acid interfaces by MALDI-MS and ESI MS/MS.

20. Wilm M, Mann M. *Anal Chem* 1996;68:1. [PubMed: 8779426]
21. Kennedy RT, Jorgenson JW. *Anal Chem* 1989;61:1128.
22. Doneanu CE, Griffin DA, Barofsky EL, Barofsky DF. *J Am Soc Mass Spectrom* 2001;12:1205. [PubMed: 11720396]
23. Lindahl T, Ljungquist S, Siebert W, Nyberg B, Sperens B. *J Biol Chem* 1977;252:3286. [PubMed: 324994]

Figure 1. Off-line HPLC Isolation of Nucleopeptides

The LC/UV trace of a nuclease P1 digest of tryptic nucleopeptides indicating the fractions labeled I and II (A). MALDI spectra of Fraction I (B) and Fraction II (C).

Figure 2. Nano ESI MS Analysis of HPLC-purified Nucleopeptides
Nanospray spectra of HPLC-purified (Figure 1A) nucleopeptides Fractions I (A) and II (B).

Figure 3. Off-line nano-LC isolation of nucleopeptides

Nano-LC/UV chromatogram produced by loading the column with 10 μ L of a nuclease P1 digest (A). The sample was eluted at 0.5 μ L/min (pressure at inline filter = 100 bar) with an exponentially produced gradient of 0% to ~ 100% solvent B (0.1 % TFA in 40% acetonitrile). UV detection was at 200 nm. Two fractions were collected manually with a fraction volume of 2 μ L (2 min collection period). One microliter of solvent A (0.1% TFA in water) was added to each vial before starting the manual collection. An aliquot (1 μ L) of each of the purified nucleopeptide samples was mixed with THAP matrix in a 1:3 ratio and analyzed by MALDI MS. MALDI spectrum exhibiting signals corresponding to $T18 \times dT_2$ (m/z 1796.1) and $T18 \times dT_3$ (m/z 2100.3) nucleopeptides isolated in nano-LC/UV Fraction 1 (B); the inset shows the isotopic distribution of $T18 \times dT_2$. MALDI spectrum of $T6_{(-5)} \times dT_2$ (m/z 2485.5) nucleopeptide isolated in nano-LC/UV Fraction 2 (C).

Figure 4. Online micro-LC/ESI MS of a Nucleopeptide Mixture

Base-peak chromatogram of a nucleopeptide mixture was produced using a micro-HPLC system coupled on-line with an ESI ion trap mass spectrometer (A). Individual mass spectra recorded at retention times of 21.5 (B), 23.6 (C), and 24.9 (D) minutes respectively.

Figure 5. Nano-LC/ESI-MS mass-chromatogram of P1 digested nucleopeptides

The chromatogram was reconstructed from m/z -signals corresponding to doubly charged molecular ions. The nucleopeptide sample (5 μ L) was injected onto the column (75 μ m i.d. $\times$ 40 cm long; 5 μ m C_{18}) and eluted at 0.3 μ L/min (pressure at inlet filter = 140 bar) with an exponentially produced gradient of 0% to ~ 100% solvent B (0.1% AA plus 0.01% TFA in 40% acetonitrile). Ion detection was from m/z 400 to 2000. All five dT_2 nucleopeptides were detected using this method.

Table I

Ung x dT20 crosslinked nucleopeptides detected by various LC/MS interfaces.

LC/MS method	Nucleopeptides				
	T6-3	T6-5	T6-7	T11	T18
Off-line microflow MALDI-TOF	✓	✓	✓	✓	✓
Off-line microflow nano ESI-MS	-	✓	✓	✓	✓
Off-line nanoflow MALDI-TOF	-	✓	-	-	✓
On line microflow ESI-MS	✓	-	-	✓	✓
On line nanoflow ESI-MS	✓	✓	✓	✓	✓