Published in final edited form as:

J Proteome Res. 2010 August 6; 9(8): 4205–4214. doi:10.1021/pr100515x.

Ultrafast Ultraviolet Photodissociation at 193 nm and its Applicability to Proteomic Workflows

James A. Madsen¹, Daniel R. Boutz², and Jennifer S. Brodbelt^{*,1}

¹Department of Chemistry and Biochemistry, The University of Texas at Austin, 1 University Station A5300, Austin, TX, USA 78712

²Center for Systems and Synthetic Biology, Institute for Cellular and Molecular Biology, The University of Texas at Austin, Austin, TX, USA 78712

Abstract

Ultraviolet photodissociation (UVPD) at 193 nm was implemented on a linear ion trap mass spectrometer for high-throughput proteomic workflows. Upon irradiation by a single 5 ns laser pulse, efficient photodissociation of tryptic peptides was achieved with production of *a*, *b*, *c*, *x*, *y*, and *z* sequence ions, in addition to immonium ions and *v* and *w* side-chain loss ions. The factors that influence the UVPD mass spectra and subsequent *in silico* database searching via SEQUEST were evaluated. Peptide sequence aromaticity and the precursor charge state were found to influence photodissociation efficiency more so than the number of amide chromophores, and the ion trap *q*-value and number of laser pulses significantly affected the number and abundances of diagnostic product ions (e.g., sequence and immonium ions). Also, photoionization background subtraction was shown to dramatically improve SEQUEST results, especially when peptide signals were low. A liquid chromatography – mass spectrometry (LC-MS) – UVPD strategy was implemented and yielded comparable or better results relative to LC-MS – collision induced dissociation (CID) for analysis of proteolyzed bovine serum albumin and lysed human HT-1080 cytosolic fibrosarcoma cells.

Introduction

Large scale, “bottom-up” characterization of cellular proteomes has been enormously successful¹ and continues to improve due to ongoing innovation in the areas of mass spectrometric instrumentation, tandem mass spectrometry (MSⁿ) techniques, sampling, separations, and bioinformatics. Typical high-throughput, bottom-up workflows consist of the separation of complex mixtures of digested proteins followed by online mass spectrometry (MS) and MSⁿ analysis. Proteins are then qualitatively and/or quantitatively identified by *in silico* database searching algorithms that interpret the MS data. Four of the most popular algorithms, SEQUEST,² MASCOT,³ OMSSA,⁴ and X! Tandem,⁵ were recently compared,⁶ and each algorithm yielded acceptable and similar results for a complex human protein sample. In terms of MS instrumentation, the introduction of the Orbitrap^{7, 8} and hybrid linear ion trap (LIT)-Orbitraps^{9, 10} have afforded significantly better mass accuracy and resolution (i.e., a resolving power of ~80,000 – 100,000 at *m/z* of around 1000)⁷, which has greatly increased selectivity in database searching for bottom-up experiments. Also, the analysis times of mass spectrometers continue to decrease. The newly released dual-pressure linear ion trap (Velos),¹¹ for example, has decreased cycle times two-

jbrodbelt@mail.utexas.edu.

Supporting Information Available: Supplementary figures and tables. This material is available free of charge via the Internet at <http://pubs.acs.org>.

fold by technological advances that eliminate prescans and allow faster scan rates in a low pressure trap. This improvement significantly increased the experimental duty cycle over more traditional trapping instruments, resulting in more protein identifications and a deeper depth of analysis into the proteome.¹¹ Moreover, the benefits and growing popularity of ultrafast separation techniques (e.g., ultra high performance liquid chromatography (UHPLC), microfluidics, capillary electrophoresis (CE), etc.), will certainly spur the development of even faster MS instrumentation as well as strategies that can efficiently sample extremely narrow peaks.

For successful protein identification, bottom-up approaches rely on the collection of informative tandem mass spectra that critically depend on the activation technique used. In recent years there have been many advances in activation technology, but there still remains no universal method that can successfully and efficiently analyze all proteins and their peptide constituents. The two most popular tandem mass spectrometric techniques used in large scale analysis of cellular proteins are the traditional collision induced dissociation (CID) and the more recently introduced electron transfer dissociation (ETD).¹² The latter has shown the most success for identifying and sequencing phosphorylated peptides and was recently shown to significantly outperform CID in characterizing the phosphoproteome of various human cells.^{13, 14} CID on the other hand, is an established, sensitive, and robust method that has successfully identified the most proteins based on analysis of complex tryptic peptide mixtures in shotgun bottom-up experiments.^{15, 16} Higher-energy C-trap dissociation (HCD), a form of CID that allows smaller m/z product ions to be detected, has recently been implemented on the newest hybrid Orbitraps and has yielded comparable results to that of CID for complex biological samples.^{10, 17} All of the techniques just described (CID, ETD, and HCD) require between 10 – 100 ms of activation time (depending on the specific instrument and analytes) for efficient precursor dissociation, posing a significant drawback with respect to implementation of high throughput strategies.

The use of single or multiple photons for activation and dissociation of peptides has also shown promise.¹⁸⁻²¹ Ultraviolet photodissociation at 193 nm of peptides was first performed in the 1980's,²²⁻²⁵ but experiments were limited to a few selected peptides on Fourier transform ion cyclotron resonance (FTICR) mass spectrometers. A recent renaissance in the use of short wavelength (157 and 193 nm)²⁶⁻³¹ ultraviolet photodissociation (UVPD) coupled mainly to time-of-flight (TOF) instruments has occurred due to the ultrafast (nanosecond) activation timescales and the rich tandem mass spectrometric information obtained (i.e., more types of backbone ions are observed compared to traditional activation methods). There have been noteworthy inroads in the implementation of MALDI-TOF-MS with UVPD for high-throughput proteomic workflows,³² although inefficient ion activation has hindered the analysis of the lowest abundance proteins/peptides,²⁷ coupled with the typical need for ~2000 spectral averages (~40 s) for each spectrum.²⁶⁻³² Another recently introduced and promising photodissociation method, femtosecond laser induced ionization/dissociation (fs-LID), produced rich fragmentation patterns for peptides (i.e., *a*, *b*, *c*, *x*, *y*, and *z* ions were all observed) using extremely fast femtosecond laser pulses, yet required rather long total activation periods (up to 200 ms).³³

If a single short UV pulse (5 ns) was employed for photodissociation for high-throughput liquid chromatography – tandem mass spectrometry (LC-MS/MS) instead of using conventional CID (30 msec activation time), an additional 1800 peaks could conceivably be analyzed per hour of analysis using traditional quadrupole ion traps.²⁰ This increase in duty cycle would be even more dramatic when using newer generation linear ion traps,¹¹ or when directly compared to strategies that use the slower activation methods such as ETD (typically 100 msec activation time). However, if the spectral information produced by UVPD pales in comparison to that afforded by gold standard methods such as CID, then the

benefits of ultrafast photoactivation will be negated. Therefore, a thorough comparison of UVPD and CID as applied to LC-MS/MS analysis and *in silico* data interpretation is warranted.

In the present study, UVPD at 193 nm is implemented on a linear ion trap mass spectrometer for high-throughput proteomic workflows. Efficient photodissociation of tryptic peptides is achieved using a single 5 ns laser pulse with minimal need for spectral averaging. The factors that affect the optimization of universal photodissociation parameters for LC-MS/MS UVPD experiments and successful *in silico* database searching via SEQUEST are investigated. This ultrafast photodissociation method yields comparable or better results compared to CID for complex samples of proteolyzed bovine serum albumin (BSA) (1 µg) and lysed human HT-1080 cytosolic fibrosarcoma cells (10 µg lysate) and represents a high-throughput LCMS/MS strategy based on photodissociation and database searching for the analysis of biologically relevant samples.

Experimental Section

Materials

All acids and buffer components, BSA, proteomics grade trypsin, and all model peptides were purchased from Sigma-Aldrich (St. Louis, MO), except the peptides KRPPGFSPFR, YRPPGFSPFR, DRVYIHPFHLVIH, DAEFRHDSGYQVHHQK, HCKFWW, GNHWAVGHLM, and ASHLGLAR, which were purchased from BACHEM (King of Prussia, PA). KAKAA and RAAAA were synthesized by Bio Basic (Ontario, Canada). Solvents for direct infusion and HPLC were obtained from Fisher Scientific (Fairlawn, NJ).

Sample Preparation

BSA was reduced by adding a 10-fold molar excess of dithiothreitol (DTT) (100 µL of 1.0 mM DTT) to 10 nmoles of BSA (10 µL of 1.0 mM protein), which was buffered to a pH of ~8.4 with ~4 µmoles ammonium bicarbonate. This solution was then incubated for one hour at 40 °C after which alkylation was performed by adding 4.0 µL of 1.0 M iodoacetamide (buffered to a pH of ~8.0 with ~20 µmoles ammonium bicarbonate). This solution was incubated at ambient temperature for 45 minutes in the dark followed by subsequent quenching by the addition of excess DTT. For enzymatic digestion, 5.0 µL of 1.0 mg/mL trypsin in 1.0 mM HCl was added to the reduced/alkylated solution followed by incubation at 37°C for 16 hours.

HT1080 cells were suspended in low-salt buffer (10 mM Tris-HCl, 10 mM KCl, 1.5 mM MgCl₂, pH 8.0) to swell, and were then lysed by dounce homogenization. The whole cell lysate was centrifuged at 1000 × g to clarify the soluble lysate and to remove the insoluble pellet. Soluble lysate (2 mg/mL protein) was denatured by addition of 2,2,2-trifluoroethanol (TFE) to 50% (vol/vol) and reduced with 15 mM DTT at 55°C for 45 min. Following reduction, the proteins were alkylated with 55 mM iodoacetamide at room temperature for 30 min. The sample was diluted in digestion buffer (50 mM Tris-HCl, 2 mM CaCl₂, pH 8.0) to a final TFE concentration of 5% (vol/vol). Trypsin was added to a 1:50 (enzyme:protein) concentration (w/w), and the sample was incubated at 37°C for 5 hrs. The digestion was quenched with 1% formic acid and sample volume was reduced to 20 µl by centrifugation under reduced pressure using a SpeedVac. Digested peptides were bound and washed on Thermo Fisher Scientific HyperSep C-18 SpinTips (San Jose, CA), resuspended in peptide buffer (95% H₂O/5% acetonitrile/0.1% formic acid) and filtered through Millipore Microcon 10 kDa centrifugal filters (Billerica, MA), with the digested peptides collected as flow-through.

Mass Spectrometry, Ultraviolet Photodissociation, and Liquid Chromatography

Analysis by mass spectrometry was performed on a Thermo Fisher Scientific LTQ XL linear ion trap (LIT) mass spectrometer (San Jose, CA) outfitted with a Coherent ExciStar XS excimer laser (Santa Clara, CA) operated at 193 nm. The laser setup was similar to that previously described,³⁴ except a CaF₂ lens was used to transmit 193 nm photons into the LIT. For UVPD experiments, the laser was pulsed once per scan with an energy of 8 mJ/pulse and a pulse duration of 5 ns. Pulse variable experiments used a laser pulse frequency of 500 Hz. For CID, a 30 ms activation time and a *q*-value of 0.25 were used. For direct infusion, peptides were diluted to 10 μ M in 50/50 MeOH/H₂O and 1% acetic acid prior to ESI-MS/MS analysis. The automated gain control (AGC) was set to 3×10^4 for MS and 1×10^4 for MSⁿ scans. An ESI voltage of 4 kV and a heated capillary temperature of 180°C were used for all mass spectrometry experiments.

Liquid chromatography was performed using a Dionex UltiMate 3000 system (Sunnyvale, CA) using a capillary flow splitter. An Agilent ZORBAX 300SB-C₁₈ column (Santa Clara, CA) (150 $\times$ 0.3 mm, 5 μ m particle size) was used for all separations. The column temperature was kept constant at 40°C. Eluent A consisted of 0.1% formic acid in water and eluent B 0.1% formic acid in acetonitrile. A linear gradient from 5% eluent B to 40% eluent B over 65 min at 5 μ L/min was used. Injections of approximately 1 μ g (20 picomoles) were used for the digested BSA sample and 10 μ g for the HT-1080 lysate sample. Data-dependent LC-MS/MS was performed in two different ways. For LC-MS/UVPD runs (i.e., only UVPD was used for activation), the first event was the full mass scan (*m/z* range of 400 – 2000) followed by ten consecutive UVPD events on the ten most abundant ions from the full mass scan with one UV pulse applied per MS/MS scan. A *q*-value of 0.1 was used for each UVPD event. For LC-MS/UVPD/CID runs (i.e., comparison between UVPD and CID), the first event was the full mass scan (*m/z* range of 400 – 2000) followed by ten alternating UVPD/CID events on the five most abundant peaks for a total of five UVPD and five CID events per cycle. A *q*-value of 0.25, an activation time of 30 ms, and a normalized collision energy (NCE) of 35% were used for each CID scan event; $NCE = (\text{voltage} \times 30) / (0.000062 \times m/z \times 0.006325)$ in which 0.000062 is the TickAmpSlope and 0.006325 is the TickAmpInt. A *q*-value of 0.1 and a default activation period of 30 μ s were used for each UVPD event. Although the actual irradiation time was only 5 ns (the duration of a single laser pulse) for UVPD, the commercial LTQ software limited the activation period to a minimum of 30 μ s. The maximum injection time for all mass scan and MS/MS events was set to 100 ms. The dynamic exclusion duration was set to 50 s, the exclusion list size allowed for 500 specified *m/z* values, and a single repeat count for LC-MS/UVPD and two repeat counts for LC-MS/UVPD/CID experiments. Each centroid mass spectrum and tandem mass spectrum was the average of three microscans. UVPD precursor dissociation efficiencies as well as pulse-variable and *q*-variable plots were calculated from ion peak areas using Origin 7.0. Photodissociation efficiency (expressed as a percentage) is defined as follows: $100 - [(\text{surviving precursor abundance} / \text{initial isolated precursor abundance}) \times 100]$, in which the precursor abundances are measured as ion peak areas.

Background Subtraction and Data Analysis

Prior to SEQUEST interpretation of the MS/MS data, UVPD mass spectra were subjected to a background subtraction procedure to reduce/eliminate photoionization products as described in detail in the results and discussion section. Background spectra were collected when no sample was being infused and thus no ions were detected in the trap, and at isolation *m/z* values of 250, 500, 750, or 1000 (isolation width of 4 *m/z*), using a protocol discussed in more detail later. Each centroid background spectra was the culmination of 20 averages of three microscans. The LC-MS/UVPD RAW files were subtracted from the background RAW files using Thermo Fisher Xcalibur version 2.0.7 software.

SEQUEST was used for *in silico* MS/MS interpretation through the Thermo Fisher Scientific Proteome Discoverer 1.0 software package. For SEQUEST, a signal:noise ratio of 3, a precursor mass tolerance of 1.5 Da, and a fragment mass tolerance of 0.8 Da were used. Product ion series for UVPD included *a*, *b*, *c*, *x*, *y*, and *z* ions, and for CID included *b* and *y* ions except where noted. UVPD and CID spectra were manually separated from LC-MS-UVPD/CID files for direct comparison. Non-redundant, bovine (25,243 sequences) and human (33,819 sequences) protein databases from the NCBI were used for searching of BSA and HT-1080 samples, respectively. Oxidation of methionines and carbamidomethyl modification of cysteines were set as dynamic and static (respectively) side chain modifications. Peptide hits were filtered against decoy databases at a 1% false discovery rate (FDR) (based on Xcorr vs. charge), and were manually verified by the presence of matching immonium ions and predicted sequence ions (from selective cleavages). Peptides with less than six amino acids were also filtered out.

Results and Discussion

Activation of multiply-charged peptide cations upon absorption of 193 nm photons produces a mixture of *a*, *b*, *c*, *x*, *y*, and *z* ions and a few *w* and *v* side-chain loss ions, which is illustrated in the collection of UVPD spectra found in this report (e.g., Supplementary Figure 1, Figure 1a, and Figure 2a). For some peptides that have a number of aromatic side-chains, occasionally some aromatic side-chain losses are also observed (labeled as “m” ions in the spectra, see Figure 2a). Foremost the *b*/*y* series followed by the *a*/*x* series are generally observed as the most abundant products, often yielding almost full sequence coverage from each of these four ion types. These spectra will be discussed in more detail later in the context of the study. As previously observed for photodissociation of peptides at 157 nm³⁰ and as also noted in this study at 193 nm (Supplementary Figure 2), singly-charged precursors result in more abundant products arising from side-chain losses as compared to those from multiply-charged precursors, which predominate in LC-MS/MS analysis. Therefore, the generation of an extensive array of abundant *a*, *b*, *c*, *x*, *y*, and *z* sequence ions in addition to immonium ions and low abundant *v* and *w* side-chain loss ions indicates that the UVPD spectra at 193 nm are well suited to *in silico* database searching, as shown in this study.

Impact of Aromatic Side-chains and Charge State on UVPD

Having a set of robust and universal activation parameters that will yield both high dissociation efficiencies and informative product ions is important for analyzing a variety of peptides of different sizes, amino acid sequences, and charge states in a single LC-MS/MS run. In the present study, the UV photodissociation efficiencies were evaluated for a series of peptides of varying lengths, charge states, and sequences, including ones containing residues with aromatic side-chains. The results are summarized in bar graph form in Figure 3a, in which the abundance of the surviving precursor is monitored relative to the abundance of the original precursor ion, thus allowing estimation of photodissociation efficiencies using a single (5 ns, 8 mJ) pulse. Figure 3a shows the photodissociation efficiencies relative to the number of aromatic side-chains, ranging from zero to three. The photodissociation efficiency increases with the number of amino acids containing aromatic side-chains (tryptophan, tyrosine, phenylalanine), with efficiencies ranging from 50% for ASHLGLAR (2+) to 98% for HCKFWW (2+). This trend was also observed for peptides with nearly identical sequences (e.g., KRPPGFSPFR, RPPGFSPFR, and YRPPGFSPFR, all 2+), but differing by one aromatic group. KRPPGFSPFR and RPPGFSPFR each contained two aromatic groups and yielded similar photodissociation efficiencies of $80 \pm 2\%$ and $78 \pm 3\%$, respectively. YRPPGFSPFR, on the other hand, contains an additional tyrosine and resulted in a dissociation efficiency of $92 \pm 2\%$.

The total abundances of the immonium product ions (based on peak areas) were plotted as a function of the number of laser pulses for the doubly-charged YRPPGFSPFR (Supplementary Figure 3). The abundances of the arginine (R), serine (S), and proline (P) immonium ions increase on going from one to two UV pulses and plateau after two laser pulses. However, for the phenylalanine (F) and tyrosine (Y) immonium ions, the abundances of the Y immonium ions steadily decrease after the first pulse and likewise the abundances of the F ions diminish after the third pulse (see inset), a result consistent with their enhanced secondary dissociation due to higher photoabsorptivity. Similarly, the decrease in the abundance of the tyrosine immonium ions substantially exceeded the decrease in abundance of the phenylalanine immonium ions, suggesting a significant difference in photoabsorptivities or stabilities of these two species. The general consensus for vacuum UVPD (157 and 193 nm) is that photodissociation efficiency is generally mediated by the photoabsorptivities of the backbone amide chromophores.^{19, 24, 31} While this is true for peptides with no F, Y, and W residues, the presence of amino acids bearing these aromatic side-chains seems to strongly influence the dissociation behavior in the present study, as noted in Figure 3a and Supplementary Figure 3, and as also noted previously for UVPD at 266 nm.³⁵

In Figure 3b, the photodissociation efficiency is shown for a series of peptides containing the same number and types of aromatic amino acids (one phenylalanine and one tyrosine), while the lengths and charge states of the peptides are varied. Each peptide from the shortest (DRVYIHPF, 8 residues) to the longest (DAEFRHDSGYQVHHQK, 16 residues) yielded nearly identical photodissociation efficiencies, ranging from 69 to 85% even when the total number of amino acids doubled. Also, as the charge state increased, the photodissociation efficiencies increased, presumably due to some combination of greater proton mobility and greater coulombic repulsion, an effect also observed previously for IR photodissociation.³⁶ The results in Figure 3 reinforce that the number of aromatic side-chains and the precursor charge state influence photodissociation efficiency more so than the number of amide chromophores.

Impact of Number of Laser Pulses and RF Trapping Voltage on UVPD

The influence of the number of laser pulses and the rf trapping voltage (as reflected by the q -value) on peptide dissociation was examined for the doubly- and triply-charged state of YRPPGFSPFR (Figure 4). The percentages of sequence, immonium, and precursor ions (based on ion peak areas) were tracked as a function of the varied parameter (i.e., number of pulses or q -value). Upon increasing the number of laser pulses while maintaining a constant q -value of 0.05, the abundances of sequence ions increased up until two laser pulses for the doubly-charged species (Figure 4a) and up until between one and two laser pulses for the triply-charged species (Figure 4b), after which the abundances of these ions decreased. As the abundances of the precursor and sequence ions decreased with increasing laser pulses, immonium ions dominated the spectra especially after the fourth pulse for the doubly-charged species and the third pulse for the triply-charged species; this trend is attributed to their formation and survival upon UVPD of the precursor ions as well as ongoing secondary dissociation of primary sequence ions into immonium ions.

Upon increasing the q -value while keeping the number of laser pulses constant (using a single 8 mJ, 5 ns pulse), higher q -values (especially increasing from 0.05 to 0.1) resulted in a greater extent of precursor dissociation and more sequence information without significant loss of low mass diagnostic ions upon activation of both the doubly- and triply-charged precursors (Figure 4c and 4d). However, at q -values above approximately 0.15, many low mass sequence and immonium ions were lost due to limitations associated with the low mass cutoff (LMCO) inherent to ion traps. Even for low m/z precursors (e.g., peptides in high charge states) where there is no significant LMCO, using higher q -values (0.1 versus 0.05)

resulted in increased sequence information as seen for triply-charged ASHLGLAR in Supplementary Figure 4. This increase in sequence information at higher q -values likely arises from the reduction in the size of the ion cloud and thus its better overlap with the laser beam in the ion trap, which has been studied in depth previously for IR photodissociation.³⁷

To afford the most informative tandem mass spectral information (with respect to sequence and immonium ions) in the most high-throughput manner, a single 5 ns (8 mJ) laser pulse and a q -value of 0.1 was used as the universal set of dissociation parameters for the remainder of the LC-MS/VPD study. Using these activation parameters, full sequence coverage from both N-terminal and C-terminal ions was achieved from a rich array of diagnostic ions for two tryptic-like peptides, YRPPGFSPFR (3+) and ASHLGLAR (2+), each which have significantly different dissociation efficiencies as previously described (see Supplementary Figure 1 for spectra). Furthermore, the presence of the side-chain loss sequence ions, w_3 and w_5 , from ASHLGLAR (Supplementary Figure 1b) allows the differentiation of the isobaric leucine and isoleucine residues; however, this ion type has yet to be incorporated into current *in silico* algorithms.

LC-MS/VPD with Background Subtraction for *In Silico* Interpretation

Upon introducing 193 nm photons into the linear ion trap, significant ion abundances are seen from photoionization products. Similar products have been observed previously using 157 nm UVPD and are thought to originate from photoionization of background organic species in the vacuum system.³⁰ This photoionization phenomenon is illustrated in Supplementary Figure 5. With no sample solution being infused or analyte ions injected into the ion trap, the region of the mass spectrum from m/z 496 to 504 was isolated to empty the trap of any residual ions (Supplementary Figure 5a). Upon subsequent irradiation of the trapping volume using one 5 ns (8 mJ) laser pulse, significant background photoionization products are observed around m/z 200 (Supplementary Figure 5b). These products are problematic because they overlap with potential low m/z diagnostic ions, impeding *in silico* spectral interpretations and reducing sensitivity. Therefore, a procedure for photoionization background subtraction was implemented and applied to subsequent LC-MS/VPD runs in conjunction with standard *in silico* database searching (via SEQUEST). The background spectrum must be representative of the types and abundances of background ions formed during UVPD spectra, thus requiring acquisition using a similar scan program (i.e., similar trapping conditions and time segments). To achieve this goal, background spectra were collected using various m/z isolation windows (e.g., 250 ± 4 , 500 ± 4 , 750 ± 4 , and 1000 ± 4) in an effort to find a universal background spectrum that could be applied to entire LC-MS/MS runs and also avoid loss of ions due to the LMCO. Photoionization abundances differed between the various m/z isolation windows used as seen in Supplementary Figure 6 due to variations in trapping efficiencies.

The effectiveness of the background subtraction on the UVPD spectra can be seen in Figure 5 for the BSA tryptic peptide LGEYGFQNALIVR (2+) using an isolation window of m/z 500 ± 4 . Supplementary Table 1 shows the SEQUEST results obtained for the tryptic digest of BSA analyzed by LC-MS/VPD both without background subtraction and with background subtraction using background spectra acquired using the four different m/z isolation windows as described above (statistics were based on three runs). For each SEQUEST score (Xcorr and probability), a higher confidence is associated with a higher score (this scoring system will be described in more detail later in this study). The use of background subtraction applied to entire LC-MS/VPD runs significantly enhanced SEQUEST identifications and scores (XCorr and probability) for all four background subtraction options as compared to using no background subtraction procedure. Background subtraction of the photoionization spectrum acquired by isolating m/z 500, 750, and 1000 all outperformed the m/z 250 isolation window. The background spectra acquired using the m/z

250 window yielded lower photoionization product abundances compared to the other three isolation windows, thus resulting in a less effective subtraction of the background ions from the UVPD spectra. From these results, background subtraction using photoionization spectra at isolation windows of m/z 500 ± 4 , 750 ± 4 , or 1000 ± 4 were applied individually to all LC-MS/VPD spectra for the remainder of the study to improve algorithm confidence and peptide identification. The best *in silico* searching results were then used for CID comparisons in subsequent sections.

LC-MS/MS: Comparing UVPD and CID

A data-dependent LC-MS/MS method was developed that enabled alternating UVPD and CID scans of each peptide ion and allowing the most direct comparison of the attributes of the resulting tandem mass spectra upon application of the search algorithms. SEQUEST comparisons of the resulting LC-MS/MS data acquired for a tryptic digest of BSA are summarized in Table 1. A major reason SEQUEST was selected as the search algorithm over other algorithms such as MASCOT and OMSSA is that all major product ion types (e.g., a , b , c , x , y , z) can be searched simultaneously in each spectrum. This capability directly benefits peptide identification and scoring for UVPD since Xcorr scores (the major peptide scoring system of SEQUEST) are a measure of how well a theoretical spectrum matches an experimental spectrum, and UVPD yields richer spectra with more a , b , c , x , y , and z ions compared to the array of ions obtained upon CID (typically b and y). In general, the UVPD spectra lead to significantly higher peptide Xcorr scores relative to those arising from the analogous CID spectra, often resulting in a higher number of peptide identifications (comparing the first column with the third column of Table 1). Examples of spectra illustrating this outcome are shown in Figure 1 for the BSA peptide RPCFSALTPDETYVPK (3+) identified via LC-MS/MS. UVPD generated significantly more sequence ions and greater coverage than did CID, the latter which suffered both from the loss of diagnostic ions from the ion trap due to the low mass cutoff and missed backbone cleavages. The resulting Xcorr scores for RPCFSALTPDETYVPK (3+) from UVPD and CID were 4.90 and 3.58, respectively.

To investigate whether this increase in scoring was a direct result of increased spectral information upon UVPD or rather an artifact of searching using a greater array of sequence ions (i.e., searching with a , b , c , x , y , z for UVPD versus b , y for CID), both the UVPD and CID spectra were searched first using a , b , c , x , y , z and then again using only b , y product ions as seen in Table 1. For the UVPD data, searching with more ions dramatically increased peptide identifications and scoring as compared to searching with only b and y ions (comparing the first and second columns of Table 1). There was no significant difference within error (statistics based on three runs), however, when searching with a , b , c , x , y , z versus just b , y for the CID data (comparing the third and fourth columns of Table 1). Therefore, the increase in scoring and number of peptide identifications obtained from the UVPD data is likely the result of more informative MS/MS spectra that contain a greater array of sequence ions than CID.

One drawback with database searching using many sequence ions is the potential for an increase in false positives even after filtering at a 1% false discovery rate (FDR). The majority of these false positives that passed the 1% FDR yielded low probability scores of 1.00, indicating a high probability that the peptide was identified by chance, and thus could be eliminated by additional filtering based on a probability score threshold. While this procedure is effective, it also eliminates many of the true positives identified by UVPD. Many of these true positives with low probability scores were a result of singly charged peptides producing many abundant side-chain loss sequence ions (e.g., v , d , w , m). For example, UVPD of the singly-charged, BSA peptide YLYEIAR (Supplementary Figure 2) yielded a good Xcorr score of 3.04, but a poor probability score of 1.00. Manual

interpretation of this spectrum showed that seven of the eleven most abundant sequence ions were *v*, *w*, and *m* side-chain loss ions with many even lower abundance side-chain loss ions also present in the spectrum. While side-chain loss ions are very useful for sequencing peptides and differentiating between leucine and isoleucine residues, they impede current *in silico* database searching algorithms, such as SEQUEST, which currently lack the ability to identify these ions. As observed herein and previously using UVPD at 157 nm,³⁰ the formation of side-chain loss ions was most notable for singly-charged peptides which are less often observed in LC-ESI-MS due to the use of acidic and aqueous eluents that usually produce peptides in charge states greater than or equal to 2+. Regardless, even after further filtering of true positive peptide hits based from low probability scores, UVPD yielded comparable and often better SEQUEST results as compared to CID.

To investigate the performance of the high-throughput-UVPD method for complex, biologically relevant samples, data-dependent LC-MS/UVPD/CID analysis was applied to a lysed human HT-1080 cytosolic fibrosarcoma cell sample. Examples of the UVPD and CID mass spectra obtained for one peptide from this facet of the study are shown in Figure 2. Figure 6 illustrates the SEQUEST results for the top twenty-five protein hits; the protein names and further SEQUEST scoring information are summarized in Supplementary Table 2. All UVPD and CID spectra for which peptides were successfully identified through SEQUEST searching were manually separated and verified for each protein. As seen in Figure 6a, both UVPD and CID yielded comparable results with respect to peptide identifications with UVPD often affording several additional peptide identifications. The sums of the peptide Xcorr scores for each protein were significantly better for UVPD as compared to CID (Figure 6b). This increase in scoring likely arises from the more informative spectra produced by UVPD versus CID as noted previously in this report for the BSA tryptic digest. For example, the UVPD mass spectrum of the doubly charged, HT-1080 peptide DLYANTVLSGGTTMYPGIADR (Figure 2a) shows many *a*, *b*, *c*, *x*, *y*, *z* product ions and yields full sequence coverage from both series of N-terminal and C-terminal ions. CID of the same peptide, however, yielded significantly less spectral information and incomplete peptide sequence coverage (Figure 2b). Thus, the Xcorr scores were 6.78 and 4.44 for UVPD and CID, respectively. No significant trend with respect to probability scores was observed for UVPD versus CID (Figure 6c); however, CID did outperform UVPD with respect to the summation of these scores for several proteins (e.g., Protein # 4, 5, 6, and 14) due to the presence of a few peptides with very high, max probability scores as seen in Supplementary Table 2. The total protein identifications were comparable between the two methods for the back-to-back LC-MS/UVPD/CID analysis. After peptide filtering at a 1% FDR followed by further filtering of peptides with probability scores less than 1.00, CID spectra identified 191 proteins and UVPD spectra identified 202 proteins. Overall, the faster UVPD method (using a single 5 ns, 8 mJ pulse) performed comparably or often better than the slower CID approach even for complex biological samples based on this high-throughput, shotgun-style proteomic strategy.

Conclusions

The results in the present study illustrate the utility of using ultrafast UVPD for high-throughput proteomic workflows. UVPD at 193 nm results in ample production of *a*, *b*, *c*, *x*, *y*, and *z* sequence ions, in addition to immonium ions and *v* and *w* side-chain loss ions. Photodissociation efficiencies range from 50-98% for doubly-charged peptides using a single laser pulse. Peptides containing one or more amino acids with aromatic side-chains exhibit higher dissociation efficiencies than peptides without aromatic groups, and the length of the peptide has relatively little impact on UVPD efficiency. A background subtraction procedure to account for the formation of rather high abundances of ions due to photoionization of background organic species during the UVPD period was implemented

and significantly improved subsequent SEQUEST scoring and peptide identifications for complex mixtures analyzed by LC-UVPD-MS. Comparable and often improved *in silico* searching results were achieved for biologically relevant samples, including a human HT-1080 cytosolic fibrosarcoma cell sample, using ultrafast UVPD in direct comparison to the gold standard method of CID.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

J.S.B acknowledges funding from the NSF (CHE-0718320), and the Welch Foundation (F-1155). D.R.B acknowledges funding from the Welch Foundation (F-1515), and the NIH (GM076536, GM067779, and GM088624).

References

1. de Godoy LMF, Olsen JV, Cox J, Nielsen ML, Hubner NC, Froehlich F, Walther TC, Mann M. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature*. 2008; 455:1251–1254. [PubMed: 18820680]
2. Eng JK, McCormack AL, Yates JR III. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom*. 1994; 5:976–89.
3. Perkins DN, Pappin DJC, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*. 1999; 20:3551–3567. [PubMed: 10612281]
4. Geer LY, Markey SP, Kowalak JA, Wagner L, Xu M, Maynard DM, Yang X, Shi W, Bryant SH. Open Mass Spectrometry Search Algorithm. *J Proteome Res*. 2004; 3:958–964. [PubMed: 15473683]
5. Craig R, Beavis RC. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics*. 2004; 20:1466–1467. [PubMed: 14976030]
6. Balgley BM, Laudeman T, Yang L, Song T, Lee CS. Comparative evaluation of tandem MS search algorithms using a target-decoy search strategy. *Mol Cell Proteomics*. 2007; 6:1599–1608. [PubMed: 17533222]
7. Hardman M, Makarov AA. Interfacing the Orbitrap Mass Analyzer to an Electrospray Ion Source. *Anal Chem*. 2003; 75:1699–1705. [PubMed: 12705605]
8. Hu Q, Noll RJ, Li H, Makarov A, Hardman M, Cooks RG. The Orbitrap: A new mass spectrometer. *J Mass Spectrom*. 2005; 40:430–443. [PubMed: 15838939]
9. Makarov A, Denisov E, Kholomeev A, Balschun W, Lange O, Strupat K, Horning S. Performance Evaluation of a Hybrid Linear Ion Trap/Orbitrap Mass Spectrometer. *Anal Chem*. 2006; 78:2113–2120. [PubMed: 16579588]
10. Olsen JV, Schwartz JC, Griep-Raming J, Nielsen ML, Damoc E, Denisov E, Lange O, Remes P, Taylor D, Splendore M, Wouters ER, Senko M, Makarov A, Mann M, Horning S. A dual pressure linear ion trap Orbitrap instrument with very high sequencing speed. *Mol Cell Proteomics*. 2009; 8:2759–2769. [PubMed: 19828875]
11. Pekar Second T, Blethrow JD, Schwartz JC, Merrihew GE, MacCoss MJ, Swaney DL, Russell JD, Coon JJ, Zabrouskov V. Dual-Pressure Linear Ion Trap Mass Spectrometer Improving the Analysis of Complex Protein Mixtures. *Anal Chem*. 2009; 81:7757–7765. [PubMed: 19689114]
12. Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A*. 2004; 101:9528–9533. [PubMed: 15210983]
13. Molina H, Horn DM, Tang N, Mathivanan S, Pandey A. Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry. *Proc Natl Acad Sci U S A*. 2007; 104:2199–2204. [PubMed: 17287340]

14. Swaney DL, Wenger CD, Thomson JA, Coon JJ. Human embryonic stem cell phosphoproteome revealed by electron transfer dissociation tandem mass spectrometry. *Proc Natl Acad Sci U S A*, Early Ed. 2009;1–6. 6 pp.
15. Molina H, Matthiesen R, Kandasamy K, Pandey A. Comprehensive Comparison of Collision Induced Dissociation and Electron Transfer Dissociation. *Anal Chem*. 2008; 80:4825–4835. [PubMed: 18540640]
16. Swaney DL, Wenger CD, Coon JJ. Value of Using Multiple Proteases for Large-Scale Mass Spectrometry-Based Proteomics. *J Proteome Res. ACS ASAP*.
17. Olsen JV, Macek B, Lange O, Makarov A, Horning S, Mann M. Higher-energy C-trap dissociation for peptide modification analysis. *Nat Methods*. 2007; 4:709–712. [PubMed: 17721543]
18. Brodbelt JS, Wilson JJ. Infrared multiphoton dissociation in quadrupole ion traps. *Mass Spectrom Rev*. 2009; 28:390–424. [PubMed: 19294735]
19. Reilly JP. Ultraviolet photofragmentation of biomolecular ions. *Mass Spectrom Rev*. 2009; 28:425–447. [PubMed: 19241462]
20. Ly T, Julian RR. Ultraviolet Photodissociation: Developments towards Applications for Mass-Spectrometry-Based Proteomics. *Angew Chem, Int Ed*. 2009; 48:7130–7137.
21. Gardner MW, Smith SI, Ledvina AR, Madsen JA, Coon JJ, Schwartz JC, Stafford GC, Brodbelt JS. Infrared Multiphoton Dissociation of Peptide Cations in a Dual Pressure Linear Ion Trap Mass Spectrometer. *Anal Chem*. 2009; 81:8109–8118. [PubMed: 19739654]
22. Bowers WD, Delbert SS, Hunter RL, McIver RT Jr. Fragmentation of oligopeptide ions using ultraviolet laser radiation and Fourier transform mass spectrometry. *J Am Chem Soc*. 1984; 106:7288–9.
23. Hunt DF, Shabanowitz J, Yates JR III. Peptide sequence analysis by laser photodissociation Fourier transform mass spectrometry. *J Chem Soc, Chem Commun*. 1987:548–50.
24. Hunt DF, Shabanowitz J, Yates JR III, Griffin PR, Zhu NZ. Tandem quadrupole Fourier-transform mass spectrometry. *Anal Chim Acta*. 1989; 225:1–10.
25. Lebrilla CB, Wang DTS, Mizoguchi TJ, McIver RT Jr. Comparison of the fragmentation produced by fast atom bombardment and photodissociation of peptides. *J Am Chem Soc*. 1989; 111:8593–8.
26. Zhang L, Reilly JP. Peptide de Novo Sequencing Using 157 nm Photodissociation in a Tandem Time-of-Flight Mass Spectrometer. *Anal Chem*. 2010; 82:898–908. [PubMed: 20058881]
27. Kim TY, Reilly JP. Time-Resolved Observation of Product Ions Generated by 157 nm Photodissociation of Singly Protonated Phosphopeptides. *J Am Soc Mass Spectrom*. 2009; 20:2334–2341. [PubMed: 19819166]
28. Shin YS, Moon JH, Kim MS. Observation of Phosphorylation Site-Specific Dissociation of Singly Protonated Phosphopeptides. *J Am Soc Mass Spectrom*. 21:53–59. [PubMed: 19836972]
29. Thompson MS, Cui W, Reilly JP. Factors that impact the vacuum ultraviolet photofragmentation of peptide ions. *J Am Soc Mass Spectrom*. 2007; 18:1439–1452. [PubMed: 17543535]
30. Kim TY, Thompson MS, Reilly JP. Peptide photodissociation at 157 nm in a linear ion trap mass spectrometer. *Rapid Commun Mass Spectrom*. 2005; 19:1657–1665. [PubMed: 15915476]
31. Morgan JW, Russell DH. Comparative studies of 193-nm photodissocn. and TOF-TOFMS anal. of bradykinin analogs: the effects of charge site(s) and fragmentation timescales. *J Am Soc Mass Spectrom*. 2006; 17:721–729. [PubMed: 16540342]
32. Zhang L, Reilly JP. De Novo Sequencing of Tryptic Peptides Derived from *Deinococcus radiodurans* Ribosomal Proteins Using 157 nm Photodissociation MALDI TOF/TOF Mass Spectrometry. *J Proteome Res. ACS ASAP*.
33. Kalcic CL, Gunaratne TC, Jones AD, Dantus M, Reid GE. Femtosecond Laser-Induced Ionization/Dissociation of Protonated Peptides. *J Am Chem Soc*. 2009; 131:940–942. [PubMed: 19128059]
34. Gardner MW, Vasicek LA, Shabbir S, Anslyn EV, Brodbelt JS. Chromogenic Cross-Linker for the Characterization of Protein Structure by Infrared Multiphoton Dissociation Mass Spectrometry. *Anal Chem*. 2008; 80:4807–4819. [PubMed: 18517224]
35. Oh JY, Moon JH, Kim MS. Chromophore effect in photodissociation at 266 nm of protonated peptides generated by matrix-assisted laser desorption ionization (MALDI). *J Mass Spectrom*. 2005; 40:899–907. [PubMed: 15934035]

36. Madsen JA, Brodbelt JS. Comparison of Infrared Multiphoton Dissociation and Collision-Induced Dissociation of Supercharged Peptides in Ion Traps. *J Am Soc Mass Spectrom.* 2009; 20:349–358. [PubMed: 19036605]
37. Remes PM, Glish GL. Mapping the Distribution of Ion Positions as a Function of Quadrupole Ion Trap Mass Spectrometer Operating Parameters to Optimize Infrared Multiphoton Dissociation. *J Phys Chem A.* 2009; 113:3447–3454. [PubMed: 19320447]

Figure 1.

(a) UVPD spectrum (193 nm) of the BSA peptide RPCFSALTPDETYVPK (3+), *q*-value of 0.1 and an activation of one 5 ns, 8 mJ pulse. **(b)** CID spectrum of BSA peptide RPCFSALTPDETYVPK (3+), *q*-value of 0.25, 30 ms activation, 35% NCE. Each spectrum is the average of three microscans. The precursor is denoted by *. The UVPD and CID spectra were collected from back-to-back LC-MS/UVPD/CID data-dependent events. The cysteine in the peptide has a carbamidomethyl modification.

Figure 2.

(a) UVPD spectrum (193 nm) of the beta actin peptide DLYANTVLSGGTTMYPGIADR (2+) from a human HT-1080 cytosolic lysed cell sample, q -value of 0.1 and an activation of one 5 ns, 8 mJ pulse. Low abundant w_4 , w_6 , and w_{14} side-chain loss ions were observed, but were not labeled on the spectrum. **(b)** CID spectrum of the beta actin peptide DLYANTVLSGGTTMYPGIADR (2+) from a human HT-1080 cytosolic cell sample, q -value of 0.25, 30 ms activation, 35% NCE. Each spectrum is the average of three microscans. The precursor is denoted by *. The UVPD and CID spectra were collected from back-to-back LC-MS/MS data-dependent events.

Figure 3.

(a) Photodissociation efficiencies of peptides (2+ charge state) with varying numbers of aromatic side-chain residues (phenylalanine, tyrosine, and tryptophan); the number of aromatic side-chains increases from left to right. (b) Photodissociation efficiencies of peptides with constant aromatic side-chain residues (one phenylalanine and one tyrosine) with varying peptide length and charge state. Each peptide was dissociated with one 5 ns, 8mJ pulse and a q value of 0.1. Photodissociation efficiency is defined as follows: $100 - [(\text{surviving precursor abundance}/\text{isolated precursor abundance}) \times 100]$. Aromatic residues are underlined.

Figure 4. Percent abundances (as measured by peak area) of diagnostic and precursor ions versus number of laser pulses for (a) [YRPPGFSPFR + 2H]²⁺ (b) [YRPPGFSPFR + 3H]³⁺. Percent abundances (as measured by peak area) of diagnostic and precursor ions versus q -value for (c) [YRPPGFSPFR + 2H]²⁺ (d) [YRPPGFSPFR + 3H]³⁺. The q -value was held at 0.045 for pulse variable experiments. For variable q -value experiments, the activation was set to a single 5 ns, 8 mJ pulse.

Figure 5.

LC-MS/MS spectrum of the BSA peptide LGEYGFQNALIVR (2+) with (a) no background subtraction and (b) with background subtraction. A q -value of 0.1 and a photoactivation of one 5 ns, 8 mJ pulse were used for each spectrum. Background subtraction was achieved by subtracting the photoionization background spectrum (photoactivation after isolation of m/z 500 ± 4) from the entire LC-MS/MS run. The precursor is denoted by *, and each spectrum is the average of three microscans.

Figure 6. SEQUEST analysis of LC-MS/VPD/CID of human HT-1080 cytosolic lysed cells: (a) number of peptides identified for VPD versus CID for each protein (b) sum of peptide Xcorr scores for each protein for VPD and CID (c) sum of peptide probability scores for each protein for VPD and CID. The top twenty-five protein hits were used for manual separation and comparison of VPD and CID data. The protein names and further scoring information are shown in Supplemental Table 2.

Table 1

SEQUEST results of LC-MS/VPD/CID of a BSA tryptic digest based on triplicate runs. VPD and CID were both searched using *a,b,c,x,y,z* and *b,y* product ions as indicated in the table. An activation of one 5 ns (8 mJ) pulse, and *q*-value of 0.1 was used for VPD. For each SEQUEST score, a higher confidence is associated with a higher score.

	VPD (<i>a,b,c,x,y,z</i>)	VPD (<i>b,y</i>)	CID (<i>b,y</i>)	CID (<i>a,b,c,x,y,z</i>)
Number of Peptides	100 ± 3	73 ± 8	80 ± 5	83 ± 3
Number of Peptides (1.1 probability filter)	80 ± 4	68 ± 7	77 ± 3	79 ± 1
Unique Peptides	37 ± 1	30 ± 1	35 ± 2	33 ± 3
Unique Peptides (1.1 probability filter)	33 ± 1	29 ± 1	34 ± 2	32 ± 3
Max Peptide Xcorr Score	7.19	5.16	5.17	5.61
Σ Peptide Xcorr Score	309 ± 16	193 ± 18	231 ± 2	231 ± 11
Max Peptide Prob. Score	300.00	126.79	129.22	126.79
Σ Peptide Prob. Score	2844 ± 187	2472 ± 174	2851 ± 59	3067 ± 117