Covalent Attachment and Dissociative Loss of Sinapinic Acid to/from Cysteine-Containing Proteins from Bacterial Cell Lysates Analyzed by MALDI-TOF-TOF Mass Spectrometry

Clifton K. Fagerquist,[a] Brandon R. Garbus,[a] Katherine E. Williams,[b] Anna H. Bates,[a] and Leslie A. Harden[a]

[a] Western Regional Research Center, Agricultural Research Service, United States Department of Agriculture, Albany, California, USA
[b] Department of Obstetrics, Gynecology, and Reproductive Sciences, University of California, San Francisco, School of Medicine, San Francisco, California, USA

We report covalent attachment via a thiol ester linkage of 3,5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid or SA) to cysteine-containing protein biomarkers from bacterial cell lysates of *E. coli* analyzed by matrix-assisted laser desorption/ionization (MALDI) mass spectrometry when using SA as the matrix. Evidence to support this conclusion is the appearance of additional peaks in the MS spectra when using SA, which are absent when using α-cyano-4-hydroxycinnamic acid (HCCA). The additional peaks appear at a mass-to-charge (m/z) $\sim$208 greater to the m/z of a more abundant protein ion peak. Protein biomarkers were identified by tandem mass spectrometry (MS/MS) using a MALDI time-of-flight/time-of-flight (TOF-TOF) mass spectrometer and top-down proteomics. Three protein biomarkers, HdeA, HdeB, and homeobox or YbgS (each containing two cysteine residues) were identified as having reactivity to SA. Non-cysteine-containing protein biomarkers showed no evidence of reactivity to SA. MS ions and MS/MS fragment ions were consistent with covalent attachment of SA via a thiol ester linkage to the side-chain of cysteine residues. MS/MS of a protein biomarker ion with a covalently attached SA revealed fragment ion peaks suggesting dissociative loss SA. We propose dissociative loss of SA is facilitated by a pentacyclic transition-state followed by proton abstraction of the β-hydrogen of the bound SA by a sulfur lone pair followed by dissociative loss of 3-(4-hydroxy-3,5-dimethoxyphenyl)prop-2-ynal. The apparent reactivity of SA to cysteine/disulfide-containing proteins may complicate identification of such proteins, however the apparent differential reactivity of SA and HCCA toward cysteine/disulfide-containing proteins may be exploited for identification of unknown cysteine-containing proteins. (J Am Soc Mass Spectrom 2010, 21, 819–832) © 2010 American Society for Mass Spectrometry

Since its development in the late 1980s [1, 2], matrix-assisted laser desorption/ionization (MALDI) has found numerous applications in the dynamic field of mass spectrometry (MS). MALDI-MS is now used in the analysis of synthetic polymers [3], in bottom-up [4] and top-down [5] proteomics, in microorganism identification [6, 7], in differential tissue imaging (healthy versus diseased) or in mapping drug distribution [8], in DNA oligonucleotide analysis [9], etc. Although MALDI has been coupled with a number of different mass analyzers (ion traps [10], FT-ICR [11], hybrid instruments [12–14], the most significant combination has been with time-of-flight (TOF) instruments [15], and more recently with tandem TOF instruments (TOF-TOF) [16, 17].

In spite of the increasing number of MALDI-MS applications, two fundamental aspects of MALDI have remained relatively constant. First, a high intensity laser pulse is used to simultaneously desorb and ionize matrix and analyte molecules, transferring them from the condensed solid phase to the gas phase. Second, the matrix must not chemically react with the analyte molecules. Chemical reactions between the matrix and analyte molecules would be highly undesirable as it would alter the chemical composition of the analyte molecules and probably their masses. Chemical reactions between the matrix and analyte molecules could render data analysis extremely complex if not impossible. Typically, MALDI matrices are weak organic acids having low vapor pressure, are solids at room temperature, and absorb at near-UV wavelengths. The most common laser wavelengths for MALDI are 337 nm from a

pulsed nitrogen laser or 355 nm from the third harmonic of a pulsed YAG laser.

MALDI-TOF-MS has found utility both as a research tool for the rapid identification of bacterial microorganisms from MS profiles of intact cells or cell extracts [6, 7, 18–31]. MALDI generated ion peaks of modest-sized protein ions collectively constitute a characteristic pattern or "fingerprint", which can be used to distinguish microorganisms at the level of genus, species, subspecies, and even strain [32–34]. Pattern recognition algorithms are used to compare spectra of unknown microorganisms to that of a spectral database of known microorganisms [35–37]. Possible limitations of pattern recognition data analysis are primarily related to variables that could affect spectral reproducibility, e.g., matrix selection, laser wavelength, ion transmission efficiency of mass analyzer, sample preparation, culture conditions, etc. Presumably, pattern recognition data analysis could also be affected by the chemical reactivity of the matrix. Such reactivity could alter the observed pattern or "fingerprint." Differences in chemical reactivity of MALDI matrices toward proteins could significantly complicate pattern recognition analysis due to the appearance of additional peaks when using one matrix versus another. This issue is particularly relevant when attempting higher taxonomic resolution, where a single protein biomarker may distinguish one strain from another closely-related strain [34].

An alternative approach to MALDI-TOF-MS analysis for microorganism identification involves utilizing information contained in bacterial genomic databases, specifically the molecular weights (MWs) of in silico putative proteins [38–43]. This approach attempts to correlate protein biomarker ion m/z to the MWs of in silico proteins derived from genomic databases. Limitations of this approach are that the microorganism (or a closely related strain) must be genomically sequenced. In addition, in silico protein MWs in genomic databases may not always correspond to the MWs of mature proteins because of post-translational modifications (PTM). In consequence, the matched proteins are not definitively identified but only tentatively assigned. Definitive protein identification would require sequence-specific information from MS/MS analysis. Matrix/protein reactivity could also complicate bioinformatic identification of a microorganism.

Fragmentation of a protein biomarker to obtain sequence-specific information represents a further extension of this bioinformatic approach to identification [44–46]. Demirev et al. definitively identified *Bacillus atrophaeus* and *Bacillus cereus*, from pure and mixed cultures, by MS/MS fragmentation of intact protein biomarkers of these microorganisms using a tandem TOF mass spectrometer [45]. Software and algorithms, developed in-house, were used identify proteins (and source microorganisms) by comparing the m/z of MS/MS fragment ions to a database of in silico fragment ions derived from bacterial protein sequences and associated MWs [45]. A similar top-down proteomic approach was used by Fagerquist et al. to identify protein biomarkers of different species of *Campylobacter* [46]. Web-based software, developed in-house, employs a simple peak matching algorithm for comparison of MS/MS fragment ions to in silico fragment ions as well as incorporating Demirev's P value algorithm for purposes of comparison. In addition, the software allowed comparisons of MS/MS fragment ions to residue-specific in silico fragment ions, e.g., D-, E-, P-specific residues, in contrast to a non-residue-specific comparison, i.e., all in silico fragment ion comparison [46].

In the current study, we have identified covalent attachment of 3,5-dimethoxy-4-hydroxycinnamic acid or sinapinic acid (SA) via a thiol ester linkage to cysteine-containing protein biomarkers from bacterial cell lysates during MALDI-TOF-TOF mass spectrometry. Covalent attachment was confirmed by the appearance of additional peaks when using SA that were not observed when ionizing by α-cyano-4-hydroxycinnamic acid (HCCA) suggesting a differential reactivity between these two matrices. The additional peaks were observed to have a m/z 208 greater than that of a prominent adjacent peak. MS/MS of these additional peaks also revealed fragment ions with and without SA attachment suggesting dissociative loss of the adduct by cleavage of the thiol ester bond. A pentacyclic transition-state and rearrangement is proposed to result in dissociative loss of the SA. The USDA web-based software was used to identify protein biomarkers with (and without) attached SA. Only cysteine-containing protein biomarkers appeared to react with SA. Three cysteine-containing protein biomarkers from a pathogenic and a nonpathogenic strain of *E. coli* were identified. Portions of this work were presented previously in an oral presentation on June 2nd 2009 at the 57th ASMS Conference on Mass Spectrometry and Allied Topics [47].

Experimental

Bacterial Cultivation

Two *E. coli* strains were used in this study: *E. coli* O157:H7 strain RM1272 (EDL933) and a non-O157:H7, nonpathogenic *E. coli* strain RM3061. *E. coli* O157:H7 strain RM1272 (EDL933) was originally isolated from raw hamburger associated with hemorrhagic colitis outbreak in 1982 and was genomically sequenced in 2001 [48]. EDL933 was generously provided by Jim Keen (USDA, Clay Center, NE). RM3061 was isolated by Dan Mills from Romaine lettuce as part of a USDA microbiological survey of produce conducted in 2002. RM3061 was analyzed using a Biolog Microstation microbial ID (Hayward, CA) and identified as a non-O157:H7 *E. coli*. RM3061 is also negative for Shiga toxin 1 (stx_1), Shiga toxin 2 (stx_{2abc}, stx_{2ex}, stx_{2f}), intimin (*eae*), hemolysin (*ehly*), and subtilase (*subA*) by polymerase chain reaction (PCR). Strains were preserved at -80 °C in Microbank beads (Pro-Lab Diagnostics, Richmond Hill, ON, Canada). The strains were propagated on LB

agar (Becton Dickinson, Sparks, MD). Inoculated agar plates were incubated overnight at 37 °C.

Bacterial Protein Extraction

Bacterial protein extraction has been described in detail previously [32–34]. In brief, bacterial cells were harvested with a 1 μL loop and transferred to screw-top conical tube containing 0.5 mL of 67% water, 33% acetonitrile, and 0.1% trifluoroacetic acid (TFA) at a pH of 2–3. Approximately 40 mg of 0.1 mm zirconia/silica beads (BioSpec Products Inc., Bartlesville, OK) were added to the tube, after which it was capped, bead-beat for 60 s and centrifuged at 10,000 rpm for 5 min.

MALDI-TOF-TOF Mass Spectrometry

MALDI-TOF-TOF analysis has been described in detail previously [46]. Briefly, equal volumes of bacterial cell lysate supernatant and saturated solution of MALDI matrix (67% water, 33% acetonitrile, and 0.1% TFA) were mixed in a snap-cap Eppendorf tube and a 0.5 μL aliquot was spotted onto a stainless steel 384-spot target. The two MALDI matrices utilized for this study were 3,5-dimethoxy-4-hydroxycinnamic acid (MW = 224.21 Da) also known as sinapinic acid (SA) and α-cyano-4-hydroxycinnamic acid (HCCA) (MW = 189.17 Da). MS data were collected in linear mode and MS/MS data were collected in reflectron-mode using a 4800 TOF-TOF proteomics analyzer (Applied Biosystems, Foster City, CA). The mass spectrometer is equipped with a pulsed solid-state YAG laser (repetition rate: 200 Hz, λ = 355 nm, pulse width: 5 ns). The instrument was externally calibrated in positive linear mode using bovine insulin (MW = 5733.58 Da), *E. coli* thioredoxin (MW = 11,673.47 Da), and horse heart apomyoglobin (MW = 16,951.55 Da), in reflectron-mode using the y-type fragment ions of glu[1]-fibrino-peptide B (MW = 1570.60) at *m/z* 175.120 and 1441.635. MS analysis was conducted in linear mode with positive ions accelerated from the source at 20 kV and traveling an effective field free length of 1.5 m before striking a multichannel plate detector (2.190 kV).

MS/MS analysis was conducted in reflectron mode with positive ions accelerated from the source at 8.0 kV. To maximize fragmentation efficiency, a higher than normal laser fluence was used. The timed ion selector (TIS) was operated at a resolution of either ±50 Da or ±100 Da. After passing the TIS, ions were decelerated to 1.70 kV before entry into a collision cell which was floated at 2.0 kV. Ions exiting the collision cell were re-accelerated to 15 kV in the second source. A precursor ion suppressor was used to deflect un-fragmented precursor ions from analysis. Fragment ions were analyzed using a two-stage reflectron mirror assembly (mirror 1: 10.910 kV, mirror 2: 18.750 kV) and detected by a multichannel plate detector (2.190 kV).

MS data were collected with 1000–2000 laser shots, whereas MS/MS data required 30,000 to 40,000 laser shots. MS and MS/MS spectra were processed using the instrument software (Data Explorer Software, version 4.9). MS spectra were processed using a noise filter (correlation factor = 0.7). MS/MS spectra were processed using an advanced baseline correction (peak width = 32, flexibility = 0.5, degree = 0.1) followed by noise removal (standard deviation = 2) followed by a Gaussian smooth (filter width = 31 points). Processed MS and MS/MS spectra were centroided, exported as an ASCII file (*m/z* versus absolute intensity), and uploaded to their respective databases in the USDA software.

Top-Down Proteomic Identification of Protein Biomarkers and Their Source Organism

The web-based software, developed in-house, for the rapid identification of protein biomarkers and their source microorganisms has been described in detail [46]. Briefly, the MS/MS fragment ions *m/z* are compared to a database of in silico fragment ions *m/z*, which are derived from bacterial protein sequences whose molecular weights are the same as the biomarker ion being analyzed. A simple peak matching algorithm and a *P* value algorithm were used to independently score/ rank identifications. Analyses were conducted involving comparison of MS/MS fragment ions to all in silico fragment ions, i.e., a non-residue-specific comparison, and/or a residue-specific comparison, e.g., D-, E-, P-specific or D-specific comparison. MS/MS fragment ions were assigned on the basis of the highest scoring identification [46].

Results and Discussion

Cysteine-Containing Protein Biomarkers from E. coli O157:H7 Strain RM1272 (EDL933)

Figure 1a shows the MS spectrum of the extracted cell lysate of *E. coli* O157:H7 strain RM1272 (EDL933) analyzed by MALDI-TOF-TOF-MS. The MS spectrum was collected in linear mode using HCCA. Figure 1b shows the MS spectrum of the same extracted cell lysate ionized using SA and analyzed by MALDI-TOF-TOF-MS. The two spectra are fairly similar, however, there are additional peaks at *m/z* 9948 and *m/z* 10,681 in Figure 1b that are not present in Figure 1a. Interestingly, the difference in *m/z* between the peak at *m/z* 9948 and the most prominent peak in the spectrum at *m/z* 9739 is Δ*m/z* 209 (or 209 Da assuming both ions represent singly charged protein ions). Similarly, the difference in *m/z* between the peak at *m/z* 10,681 and the second most abundant peak in the spectrum at *m/z* 10,475 is Δ*m/z* 206. There may be greater error in Δ*m/z* due to the lower intensity of the peak at *m/z* 10,681.

Figure 2a shows the MS/MS spectrum of the protein biomarker ion at *m/z* 9738 shown in Figure 1a. This protein was identified by top-down proteomics as the acid stress chaperone-like protein HdeA [49]. Prominent fragment ions are identified by their *m/z* and
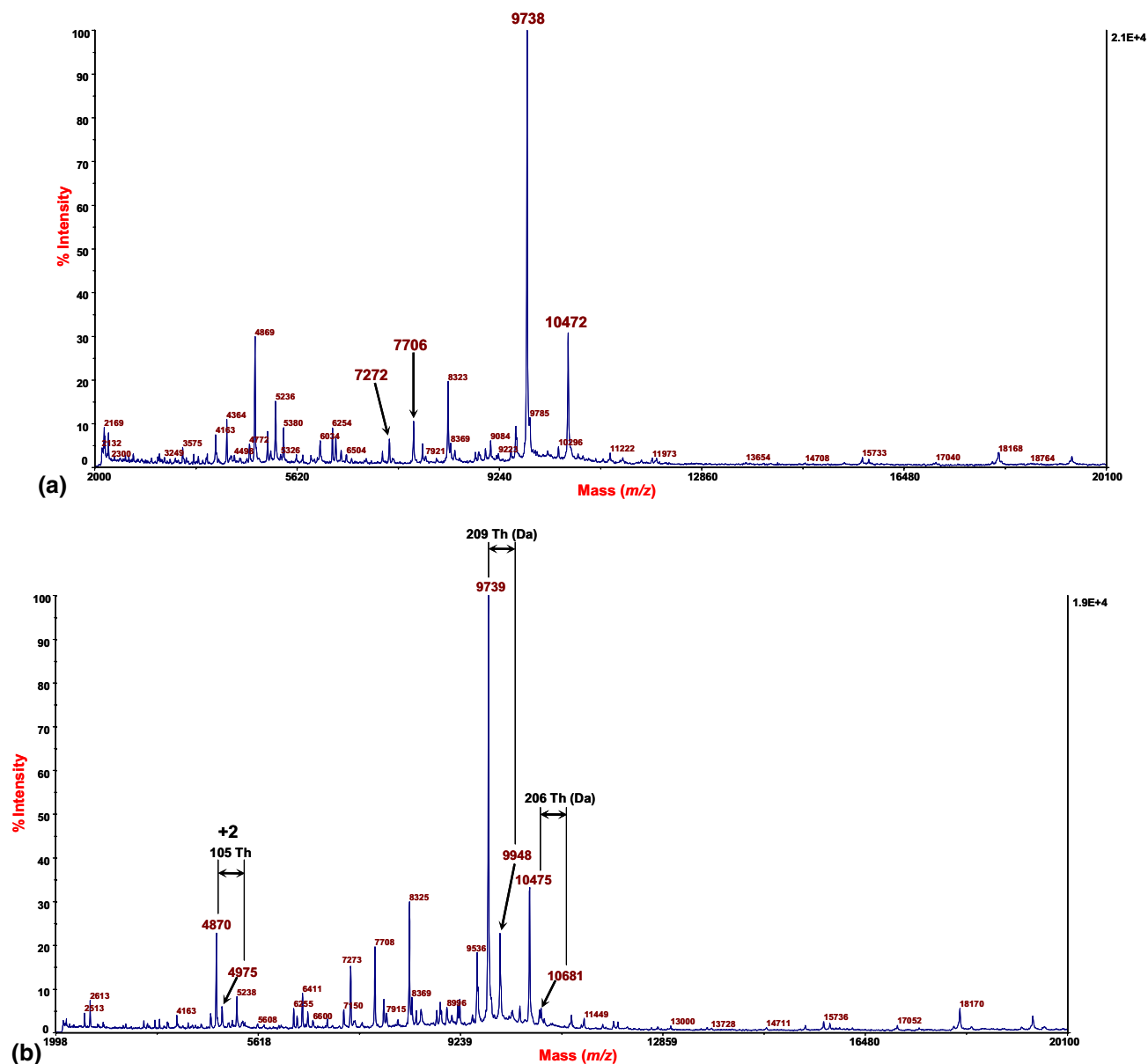
Figure 1.   (**a**) MS spectrum of the extracted cell lysate of *E. coli* O157:H7 strain RM1272 (EDL933) using HCCA matrix. (**b**) MS spectrum of the extracted cell lysate of *E. coli* O157:H7 strain RM1272 (EDL933) using SA matrix. Additional peaks appear at *m/z* 9948 and 10,681 compared with (**a**).

type/number and amino acid residues adjacent to the site of polypeptide backbone cleavage of the HdeA sequence. As can be seen, many of the fragment ions are the result of polypeptide cleavage adjacent to an aspartic acid (D) or glutamic acid (E) residue. HdeA has a 21-residue, N-terminal signal peptide and two cysteine residues suggesting a disulfide bridge. A disulfide bond is also suggested from the y_{72}-17 fragment ion at *m/z* 7897 in which the site of polypeptide backbone fragmentation is flanked by a threonine (T) and cysteine (C) residue. A disulfide bridge involving Cys18 and Cys66 residues, in the mature protein, may facilitate protein ion fragmentation by limiting efficient energy redistribution by constraining ro-vibrational degrees-of-freedom in the molecular ion.

Figure 2b shows the MS/MS spectrum of the protein biomarker ion at *m/z* 9948 shown in Figure 1b. The MS/MS spectrum is very similar to the MS/MS spectrum in Figure 2a. Several prominent fragment ions appear with approximately the same *m/z* as that of fragment ions observed in Figure 2a. In addition, fragment ions also appear with a *m/z* approximately *m/z* 206 greater than the *m/z* of an adjacent ion. The strong similarity between Figure 2a and b suggest significant similarity between the two protein ion sequences even though the *m/z* of the two precursor ions differ by ~*m/z* 208. Top-down proteomics analysis identified this protein biomarker ion as HdeA even though the *m/z* of the protein biomarker was over 200 Da greater than the molecular weight (MW) of HdeA [49]. The sudden appearance of the
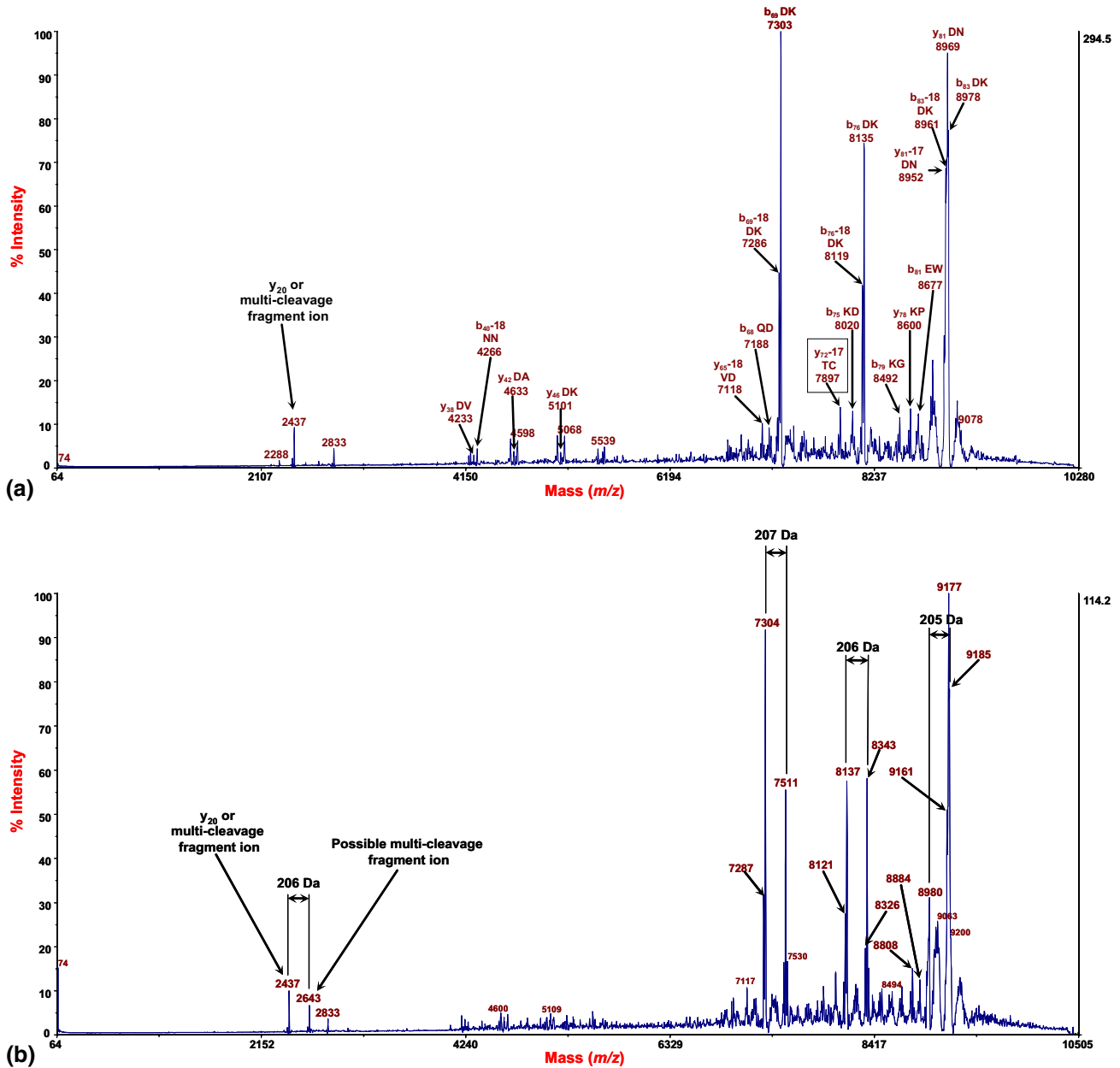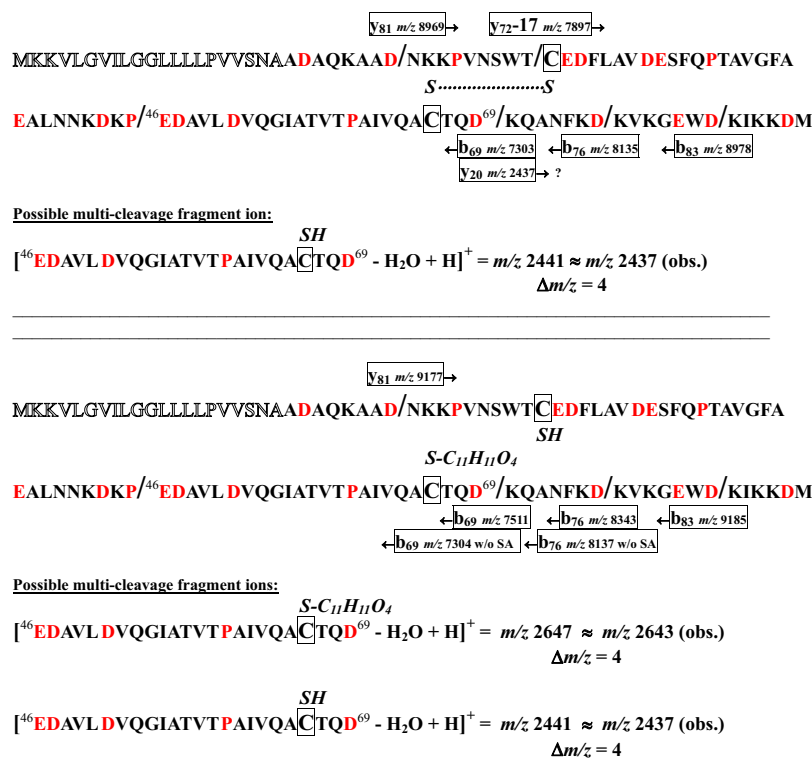
Figure 2.   (a) MS/MS spectrum of the protein biomarker ion $[M + H]^+$ at m/z 9738 (Figure 1A) from *E. coli* O157:H7 strain RM1272 (EDL933) ionized with HCCA matrix and identified by top-down proteomics as the acid stress chaperone-like protein HdeA [49]. Prominent fragment ions are identified by their m/z, type/number, and amino acid residues adjacent to the site of polypeptide backbone cleavage. The y_{72}-17 fragment ion at m/z 7897 (boxed) suggests the presence of a disulfide bridge between the two cysteine residues (Scheme 1, top panel). (b) MS/MS spectrum of the protein biomarker ion $[M + H]^+$ at m/z 9948 (Figure 1B) from *E. coli* O157:H7 strain RM1272 (EDL933) ionized with SA matrix and identified by top-down proteomics as the acid stress chaperone-like protein HdeA with an increased mass ~208 Da [49].

peak at m/z 9948 when ionizing with SA suggested covalent attachment of SA to the protein. SA has a MW of 224.2 Da, however a covalent thiol ester linkage between SA and one of the two cysteine residues of HdeA would increase the mass of the protein by ~208 Da [50–54].

The top panel of Scheme 1 rationalizes the cleavage of the polypeptide backbone of HdeA to generate the fragment ions observed in Figure 2a. As can be observed, the four most abundant fragment ions in Figure

2a are generated by cleavage of the polypeptide backbone adjacent to an aspartic acid (D) residue generating three b-type fragment ions and one y-type fragment ion. The fragment ion at m/z 2437 in Figure 2a would appear to correspond exactly to a y_{20} fragment ion, however another possible assignment is that the fragment ion at m/z 2437 may result from a b-type cleavage and y-type cleavage between residues Pro45/Glu46 and Asp69/Lys70 in the mature protein. Although the difference

$\boxed{\text{y}_{81} \ m/z \ 8969} \rightarrow$ $\boxed{\text{y}_{72}\text{-}17 \ m/z \ 7897} \rightarrow$

MKKVLGVILGGLLLLPVVSNAA**D**AQKAA**D**/NKK**P**VNSWT/$\boxed{\text{C}}$**E**DFLAV **DE**SF**Q**PTAVGFA

$S \cdots\cdots\cdots\cdots\cdots S$

EALNNK**D**K**P**/^{46}**ED**AVL **D**VQGIATVT**P**AIVQA$\boxed{\text{C}}$TQ**D**^{69}/KQANFK**D**/KVKG**E**W**D**/KIKK**D**M

$\leftarrow \boxed{\text{b}_{69} \ m/z \ 7303}$ $\leftarrow \boxed{\text{b}_{76} \ m/z \ 8135}$ $\leftarrow \boxed{\text{b}_{83} \ m/z \ 8978}$

$\boxed{\text{y}_{20} \ m/z \ 2437} \rightarrow$?

Possible multi-cleavage fragment ion:

$$SH$$
$$[^{46}\textbf{ED}\text{AVL }\textbf{D}\text{VQGIATVT}\textbf{P}\text{AIVQA}\boxed{\text{C}}\text{TQ}\textbf{D}^{69} - H_2O + H]^+ = m/z \ 2441 \approx m/z \ 2437 \ (\text{obs.})$$
$$\Delta m/z = 4$$

_____

_____

$\boxed{\text{y}_{81} \ m/z \ 9177} \rightarrow$

MKKVLGVILGGLLLLPVVSNAA**D**AQKAA**D**/NKK**P**VNSWT$\boxed{\text{C}}$**E**DFLAV **DE**SF**Q**PTAVGFA

$$SH$$

$$S\text{-}C_{11}H_{11}O_4$$

EALNNK**D**K**P**/^{46}**ED**AVL **D**VQGIATVT**P**AIVQA$\boxed{\text{C}}$TQ**D**^{69}/KQANFK**D**/KVKG**E**W**D**/KIKK**D**M

$\leftarrow \boxed{\text{b}_{69} \ m/z \ 7511}$ $\leftarrow \boxed{\text{b}_{76} \ m/z \ 8343}$ $\leftarrow \boxed{\text{b}_{83} \ m/z \ 9185}$

$\leftarrow \boxed{\text{b}_{69} \ m/z \ 7304 \ \text{w/o SA}}$ $\leftarrow \boxed{\text{b}_{76} \ m/z \ 8137 \ \text{w/o SA}}$

Possible multi-cleavage fragment ions:

$$S\text{-}C_{11}H_{11}O_4$$
$$[^{46}\textbf{ED}\text{AVL }\textbf{D}\text{VQGIATVT}\textbf{P}\text{AIVQA}\boxed{\text{C}}\text{TQ}\textbf{D}^{69} - H_2O + H]^+ = m/z \ 2647 \approx m/z \ 2643 \ (\text{obs.})$$
$$\Delta m/z = 4$$

$$SH$$
$$[^{46}\textbf{ED}\text{AVL }\textbf{D}\text{VQGIATVT}\textbf{P}\text{AIVQA}\boxed{\text{C}}\text{TQ}\textbf{D}^{69} - H_2O + H]^+ = m/z \ 2441 \approx m/z \ 2437 \ (\text{obs.})$$
$$\Delta m/z = 4$$

Scheme 1. Top panel: Amino acid sequence of the acid stress chaperone-like protein HdeA of *E. coli* O157:H7 strain EDL933. The mature protein is post-translationally modified with removal of a 21-residue N-terminal signal peptide (in outline) and a disulfide bridge (*S . . . S*) between the two cysteine residues (boxed). *SH* symbolizes a sulfhydryl group. The 89-residue mature protein average MW = 9738.91 Da. Forward slashes indicate sites of polypeptide cleavage and resulting fragment ions observed in Figure 2A. Bottom panel: Amino acid sequence of the acid stress chaperone-like protein HdeA of *E. coli* O157:H7 strain EDL933 with attached SA covalently bound via thiol ester linkage to one of the two cysteine residues. $S\text{-}C_{11}H_{11}O_4$ and *SH* symbolizes SA attachment at cysteine residues. The average MW of the 89-residue mature protein with SA covalent bound adduct is 9947.12 Da. Forward slashes indicate sites of polypeptide cleavage and resulting fragment ions observed in Figure 2b. The difference in *m/z* between a fragment ion identified in the bottom panel and its corresponding fragment ion in the top panel is $\Delta m/z$ ~208.

between the theoretical and observed *m/z* for this fragment ion is $\Delta m/z$ 4, the possibility of multi-cleavage fragment ions is supported by the observation of fragment ion at *m/z* 2643 in Figure 2b which is *m/z* 206 higher than an adjacent ion at *m/z* 2437. The additional fragment ions ~*m/z* 206 higher than adjacent fragment ions in Figure 2b strongly suggests dissociative loss of SA via a molecular ion rearrangement followed by cleavage of thiol ester linkage from the cysteine residue. In consequence, the assignment of the fragment ions at *m/z* 2437 (Figure 2a) and *m/z* 2437 (Figure 2b) as being y_{20} is probably not correct. The fragment ion $\Delta m/z$ 4 error may be due to drift in the TOF calibration affecting the lower *m/z* range [46, 55]. However, a more probable explanation is that fragment ions, generated from multiple fragmentation events, may not have attained full ion acceleration depending on the sequence of fragmentation events relative to the pulsed acceleration/de-acceleration/re-acceleration events. We have previously reported observation of multiple cleavages of the polypeptide backbone of singly charged (protonated) proteins using MALDI-TOF-

TOF-MS [56, 57]. For instance, Fagerquist et al. reported multiple fragmentations of the ~13 kDa ribosomal L7/L12 protein from different *Campylobacter*. These multi-cleavage fragment ions were identified, not only from their *m/z*, but also from known amino acid substitutions in the protein sequence of fragment ions across species/strains of *Campylobacter* [56, 57].

The bottom panel of Scheme **1** rationalizes the fragment ions observed in Figure 2b. The difference in *m/z* between a fragment ion identified in the bottom panel and its corresponding fragment ion in the top panel is ~$\Delta m/z$ 208. For example, subtracting the *m/z* of the y_{81} fragment ion at *m/z* 8969 (top panel) from the *m/z* of the y_{81} fragment ion at *m/z* 9177 (bottom panel) equals $\Delta m/z$ 208. These fragment ion comparisons are consistent with SA attachment at the two cysteine residues via a thiol ester linkage. We propose that prominent fragment ions observed in Figure 2b, which are also observed in Figure 2a, are the result of dissociative loss of SA from a cysteine-containing fragment ion.

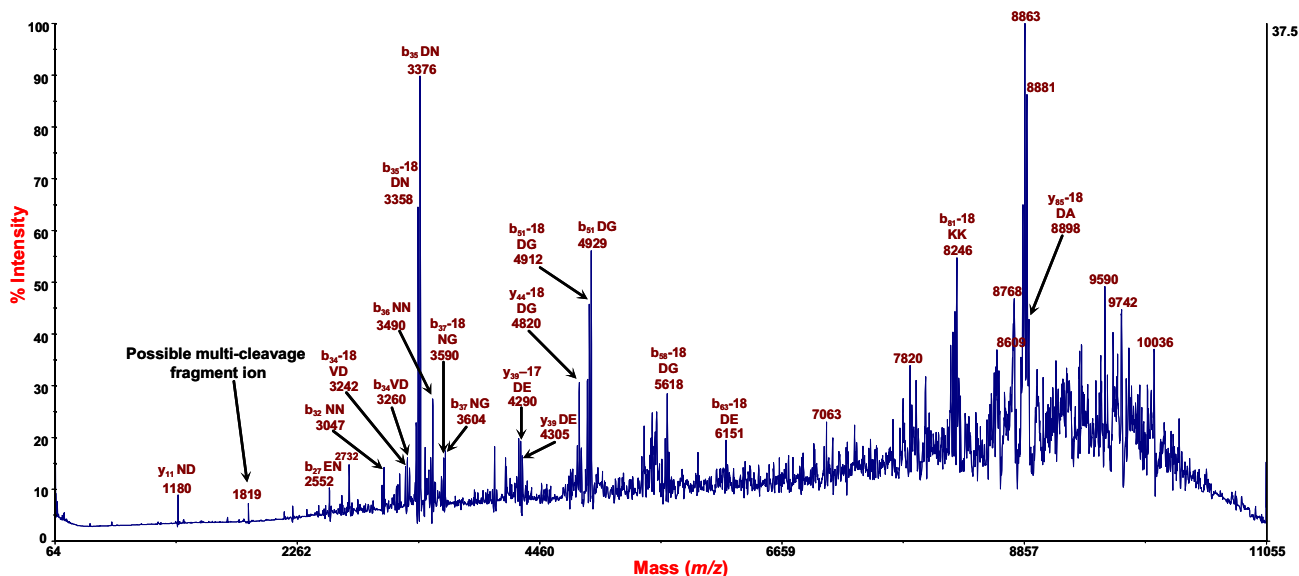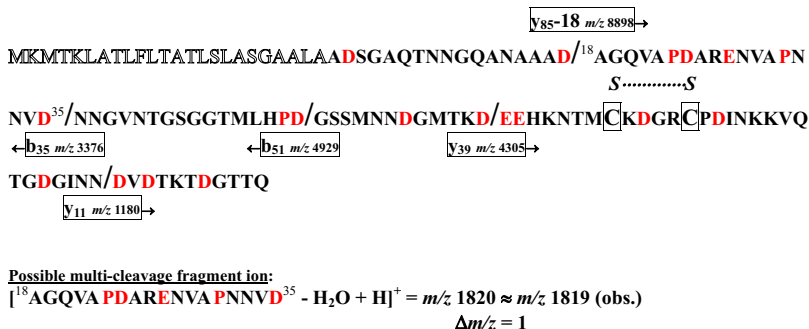Figure 3 shows the MS/MS spectrum of the protein biomarker ion at *m/z* 10,472 shown in Figure 1a. This

Figure 3. MS/MS spectrum of the protein biomarker ion $[M + H]^+$ at m/z 10,472 (Figure 1a) from *E. coli* O157:H7 strain RM1272 (EDL933) ionized with HCCA matrix and identified by top-down proteomics as the putative homeobox protein [49].

biomarker was identified by top-down proteomics as the putative homeobox protein *E. coli* O157:H7 strain EDL933 [49]. The homeobox protein has a 24-residue N-terminal signal peptide as well as two cysteine residues suggesting a disulfide bridge. The two cysteine residues are consistent with the appearance of an additional peak at m/z 10,681 in Figure 1b when ionizing with SA. A thiol ester linkage between SA and one of the cysteine residues of the homeobox protein would produce an ion with a ~m/z 208 greater than the m/z of the homeobox protein ion. As noted earlier, the m/z difference is $\Delta m/z$ 206, but this is likely due to the relatively low signal intensity of the peak at m/z 10,681. The low intensity of the ion at m/z 10,681 made MS/MS not feasible. Scheme 2 rationalizes the cleavage of the 102-residue polypeptide chain of the mature homeobox protein. The fragment ion at m/z 1819 in Figure 3 does

not correspond to any single-cleavage in silico fragment ions, however, it may correspond to a multi-cleavage fragment ion involving both a b-type and y-type cleavages between residues Asp17/Ala18 and Asp35/Asn36. The discrepancy between the theoretical and observed m/z is only $\Delta m/z$ 1 which, as mentioned previously, may be due to drift in the TOF calibration or the result of multi-cleavage fragment ions not attaining full acceleration.

Cysteine-Containing Protein Biomarker from Non-O157:H7, Nonpathogenic E. coli

Figure 4a shows the MS spectrum of the extracted cell lysate of nonpathogenic/non-O157:H7 *E. coli* strain RM3061 obtained by MALDI-TOF-TOF-MS in linear mode using HCCA matrix. Figure 4b shows the same

Scheme 2. Amino acid sequence of the putative homeobox protein *E. coli* O157:H7 strain EDL933. The mature protein is post-translationally modified with removal of a 24-residue N-terminal signal peptide (in outline) and a disulfide bridge ($S \ldots S$) between the two cysteine residues (boxed). Average MW of 102-residue mature protein = 10,473.06 Da. Forward slashes indicate sites of polypeptide cleavage and resulting fragment ions observed in Figure 3. Note the fragment ion at m/z 1819.9 may correspond to a b-type *and* a y-type cleavage of the polypeptide backbone.
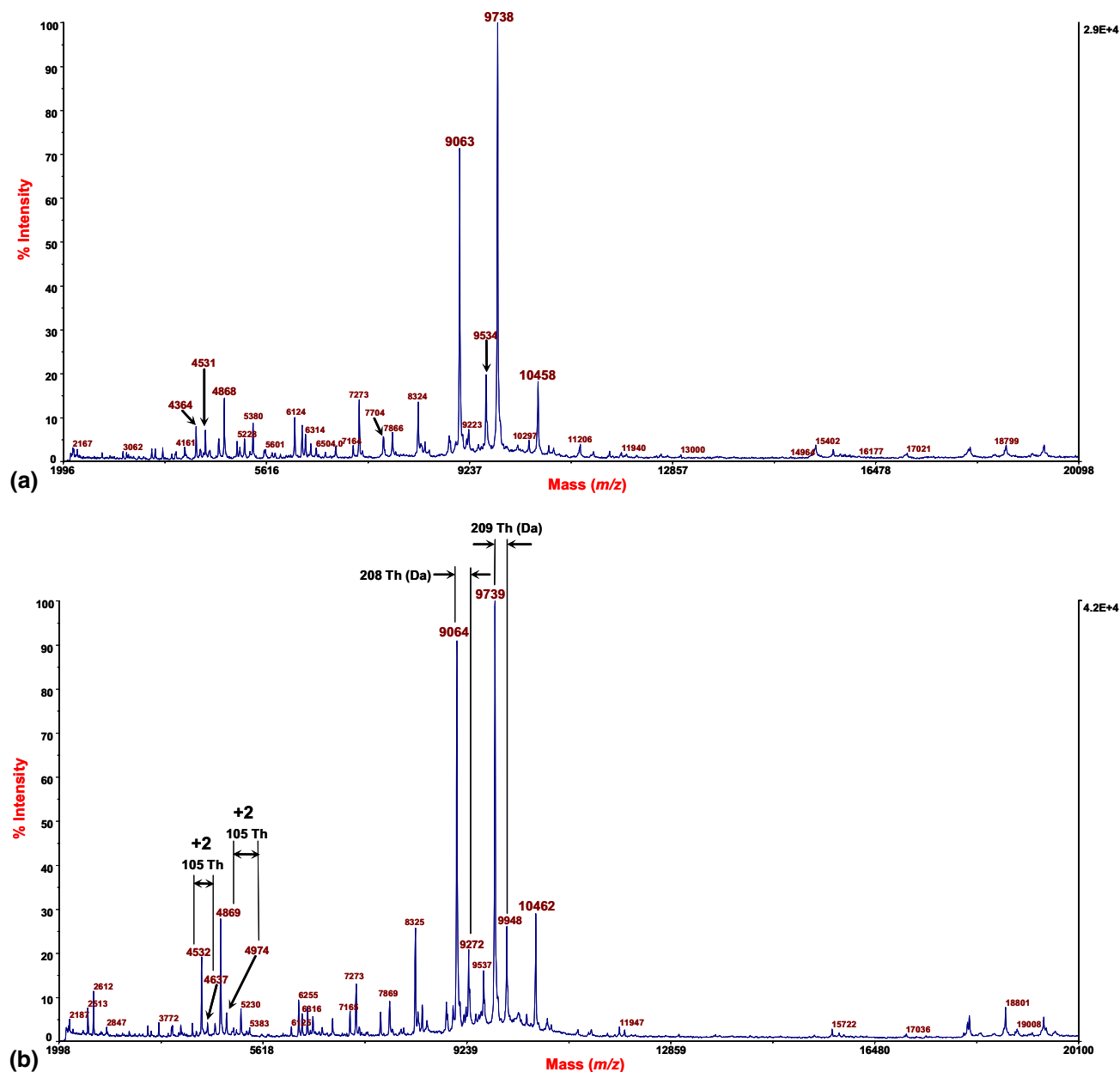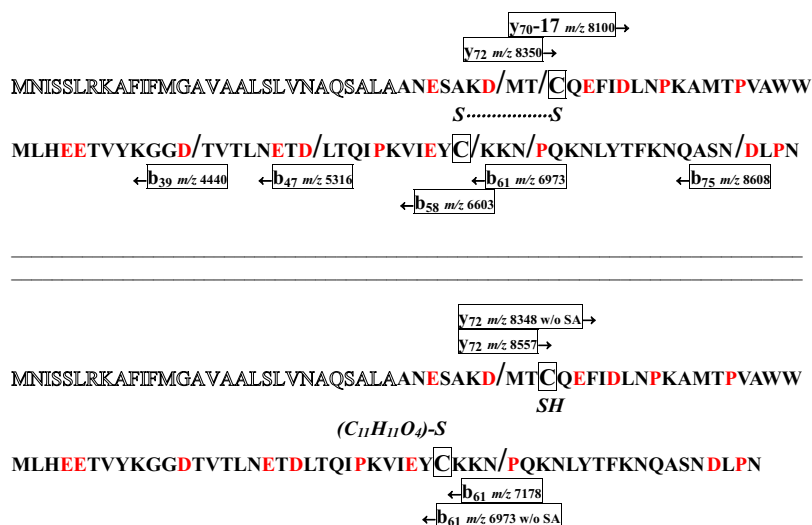
Figure 4. (**a**) MS spectrum of the extracted cell lysate of a non-O157:H7, nonpathogenic *E. coli* strain RM3061 using HCCA matrix. Note the protein biomarker ion at *m/z* 9063 which is absent from the MS spectra of *E. coli* O157:H7 strains. (**b**) MS spectrum of the extracted cell lysate of non-O157:H7, nonpathogenic *E. coli* strain RM3061 using SA matrix. Additional peaks appear at *m/z* 9272 and 9948 compared with (**a**).

analysis using SA matrix. The two MS spectra are very similar, however, there are additional peaks at *m/z* 9272 and 9948 in Figure 4b that are not present in Figure 4a. The peaks at *m/z* 9272 and 9948 are ~208 greater in *m/z* than the first and second most abundant peaks at *m/z* 9064 and 9739, respectively. Figure 5a shows the MS/MS spectrum of the protein biomarker ion at *m/z* 9064 shown in Figure 4b. This protein was identified by top-down proteomics as the acid stress chaperone-like protein HdeB [49]. HdeB has a 29-residue N-terminal signal peptide and two cysteine residues. A disulfide bridge between the two cysteines is suggested from two

fragment ions (boxed) in Figure 5a. The HdeB ion is conspicuously absent from MS spectra of the *E. coli* O157:H7 strain shown in Figure 1a and b. Mandrell et al. reported previously that the absence of the HdeB ions in MALDI-TOF-MS spectra of *E. coli* O157:H7 strains was due to mis-sense mutation in the start codon of *hdeB* that causes this protein to not be expressed [58]. In consequence, HdeB could serve as a "negative" biomarker in that it has been shown to be reproducibly absent from MS spectra of *E. coli* O157:H7 strains in this study and previous studies [49, 58, 59]. The top panel of Scheme **3** rationalizes fragmentation of the mature

Figure 5. (a) MS/MS spectrum of the protein biomarker ion $[M + H]^+$ at m/z 9064 (shown in Figure 4B) from non-O157:H7, nonpathogenic *E. coli* strain RM3061 ionized with SA matrix and identified by top-down proteomics as the acid stress chaperone-like protein HdeB [49]. Note the fragment ions (boxed) adjacent to cysteine residues. (b) MS/MS spectrum of the protein biomarker ion $[M + H]^+$ at m/z 9272 (shown in Figure 4b) from non-O157:H7, nonpathogenic *E. coli* strain RM3061 ionized with SA matrix.

HdeB as cleavage of the polypeptide backbone with the most abundant fragment ions generated from polypeptide cleavage at sites adjacent to aspartic acid (D) and glutamic acid (E) residues. There are also fragment ions generated by cleavage at sites adjacent to the two cysteine residues suggesting the presence of a disulfide bridge in the mature protein. Figure 5b shows the MS/MS spectrum of the protein biomarker ion at m/z 9272 shown in Figure 4b. The quality of the MS/MS spectrum was poorer than that shown in Figure 5a,

and it was not possible to identify the protein by top-down proteomic identification using our in-house developed software, however, a few of the more prominent fragment ions are common to the MS/MS spectrum in Figure 5a. The bottom panel of Scheme 3 rationalizes fragmentation of the HdeB with a covalently attached SA and corresponding fragment ions from Figure 5b.

The ion at m/z 9739 in Figure 4b was analyzed by MS/MS and top-down proteomics and identified as the

y_{70}-17 *m/z* 8100 →

y_{72} *m/z* 8350 →

MNISSLRKAFIFMGAVAALSLVNAQSALAANESAKD/MT/C QEFIDLNPKAMTPVAWW

S ··············· S

MLHEETVYKGGD/TVTLNETD/LTQIPKVIEYC/KKN/PQKNLYTFKNQASN/DLPN

← b_{39} *m/z* 4440   ← b_{47} *m/z* 5316   ← b_{61} *m/z* 6973   ← b_{75} *m/z* 8608

← b_{58} *m/z* 6603

_____

_____

y_{72} *m/z* 8348 w/o SA →

y_{72} *m/z* 8557 →

MNISSLRKAFIFMGAVAALSLVNAQSALAANESAKD/MTC QEFIDLNPKAMTPVAWW

SH

$(C_{11}H_{11}O_4)$-S

MLHEETVYKGGDTVTLNETDLTQIPKVIEYCKKN/PQKNLYTFKNQASNDLPN

← b_{61} *m/z* 7178

← b_{61} *m/z* 6973 w/o SA

Scheme 3. Top panel: Amino acid sequence of the acid stress chaperone-like protein HdeB of non-O157:H7, nonpathogenic *E. coli* strains RM3061 and K-12. The mature protein is post-translationally modified with removal of a 29-residue N-terminal signal peptide (in outline) and a disulfide bridge (*S . . . S*) between the two cysteine residues (boxed). Average MW of 79-residue mature protein = 9063.26 Da. Forward slashes indicate sites of polypeptide cleavage and resulting fragment ions observed in Figure 5a. Bottom panel: Amino acid sequence of the acid stress chaperone-like protein HdeB of *E. coli* strains RM3061 and K-12 with SA covalently bound via thiol ester linkage to one of the two cysteine residues (boxed). *S*-$(C_{11}H_{11}O_4)$ and *SH* symbolizes adduct attachment at cysteine residues. The average MW of the 79-residue mature protein with SA adduct = 9271.47 Da. Forward slashes indicate sites of polypeptide cleavage and resulting fragment ions observed in Figure 5b.

acid stress chaperone-like protein HdeA of *E. coli* strain K-12 [49]. As noted before, the amino acid sequence of this protein is identical to HdeA of *E. coli* O157:H7 strain EDL933 as well as *Shigella flexneri*. The ion at *m/z* 10,462 in Figure 4b was also analyzed by MS/MS and top-down proteomics and identified as the uncharacterized protein YbgS of *E. coli* strain K-12 [49]. Like the homeobox protein of *E. coli* strain EDL933, the YbgS of *E. coli* strain K-12 has a 24-residue N-terminal signal peptide and two cysteine residues suggesting a disulfide bridge. However, the two sequences differ by a single amino acid substitution, Pro50 ↔ Ser50, in the mature protein resulting in a protein MW difference of 10 Da [49].

Reactivity of MALDI Matrices Toward Cysteine-Containing Proteins

Our group was the first to report the covalent attachment via a thiol ester linkage of SA to cysteine-containing proteins from bacterial cell lysates using standard MALDI sample preparation methods [47]. In the same report, we noted a lack of reactivity of HCCA to the same cysteine-containing proteins. Possibly, the α-cyano group of HCCA (or the absence of an α-hydrogen) may inhibit (or prohibit) formation of a thiol ester linkage to cysteine residues. Cysteine residues play a critical role in the structure and function of proteins. Their primary function is the formation of disulfide bridges, which provide secondary and tertiary structure crucial to protein function. Detection of cysteine resi-

dues and/or disulfide bonds in unknown proteins during a MALDI-MS analysis is potentially useful tool for their identification. The apparent differential reactivity of SA and HCCA toward cysteine residues suggests that this difference could be exploited to identify unknown cysteine-containing protein biomarkers analyzed by MALDI mass spectrometry [47]. Only protein biomarkers having cysteine residues and/or disulfide bonds appeared to chemically react with SA. A chemical analogue of SA is 4-hydroxy-cinnamic acid (or *p*-coumaric acid), which has been identified as the photoactive co-factor of the photoactive yellow protein (PYP) from *Ectothiorhodospira halophila* [50–54]. P-coumaric acid is covalently bound to PYP at Cys69 by a thiol ester bond [50–54]. Condensed phase spectroscopy absorbance experiments of PYP using a series of chemical/chromophoric analogues of p-coumaric acid, including SA and ferulic acid (FA), were conducted that concluded that the covalent thiol ester linkage of the chromophore to the Cys69 was critical to the photoactive function of the PYP protein [50–53]. In addition, it is believed that the thiol ester-linked chromophore may undergo *cis-trans* isomerization of the *trans* vinyl bond as part of the photocycle in solution [53].

Very recently, Yang et al. reported reducing disulfide-containing peptides under alkaline conditions and reacting them with a series of MALDI matrices to generate covalently bound peptide/matrix adducts [60]. Peptide/matrix complexes were only observed with reduced peptides. Other researchers had previously noted HCCA attachment under disulfide-reducing

conditions [61, 62]. In these reports, the peptide/matrix complexes have m/z that are greater than the m/z of the reduced peptide by an amount equivalent to integer additions of the matrix MW up to the number cysteine residues present in the peptide. Yang et al. proposed a Michael addition reaction involving nucleophilic attack by a deprotonated cysteine thiol at the β-carbon of the MALDI matrix facilitated by the electron withdrawing of the α-cyano group of α-cyano-4-hydroxycinnamic and α-cyano-3-hydroxycinnamic matrices. Other α,β-unsaturated carboxylic acid matrices (e.g., sinapinic acid, caffeic acid, and ferulic acid) were also found to react but with less efficiency due, presumably, to the lack of an α-cyano group [60]. Yang et al. did not appear to observe (and did not report) covalent attachment of MALDI matrices to cysteine-containing peptides via a thiol ester linkage. Our bacterial cell lysate samples were prepared for MALDI analysis in the absence of a reducing agent and at low pH (2–3) conditions very different from that utilized by Yang et al. and others [60–62]. The formation of a thiol ester linkage between sinapinic acid and cysteine/disulfide-containing proteins from bacterial cell lysates is likely to be an acid catalyzed mechanism and thus different from the mechanism proposed by Yang et al. [60]. This is not surprising, as the matrix attachment is different by mass of 16 Da. The α-cyano group in HCCA (or the absence of an α-hydrogen) would appear to inhibit (or prohibit) thiol ester bond formation under acidic conditions. In contrast, the same α-cyano group would appear to facilitate the reaction mechanism proposed by Yang et al. under disulfide-reducing, alkaline conditions. In consequence, the relative reactivity of MALDI matrices to cysteine/disulfide-containing peptides/proteins by these two different reactions appears to be reversed.

MALDI Matrix Reactivity and Pattern Recognition Algorithms for Microbial Identification

Chemical reactivity of MALDI matrices toward bacterial proteins represents an additional experimental variable that is likely to complicate identification of microorganisms using pattern recognition algorithms [35–37]. It has already been shown in numerous experiments that the relative intensity of protein biomarker ions may change depending on the matrix employed. For instance, HCCA often results in an increase in the relative intensity of doubly charged ions compared with singly charged ions as well as difficulty in ionizing higher molecular weight proteins. In contrast, SA and/or FA often demonstrate better ionization of higher molecular weight proteins, and lower relative intensity of doubly charged ions. The fact that SA can also apparently react with cysteine residues and/or disulfide bridges resulting in the appearance of additional protein peaks (both singly and doubly charged), which are not observed with other matrices, will likely complicate multi-laboratory comparison of data for several reasons [37].

First, other MALDI matrices may also react with proteins under various sample preparation conditions. The current study was not an exhaustive study of MALDI matrix reactivity, however in re-examining previously collected data from our laboratory, we found some evidence (data not shown) suggesting that FA may also react with cysteine-containing proteins although perhaps not as efficiently as SA. Second, it is not clear, now, whether the efficiency of the cysteine residue/SA reaction and is pH-dependent or solvent-dependent. In the current study, equal volumes of bacterial supernatant and saturated matrix solution were mixed in the presence of 0.1% TFA prior to spotting on the target. However, many MALDI practitioners spot the sample on the target allowing it dry before spotting the matrix on top. Others reverse the order and deposit matrix first, and still others employ a "sandwich" technique with the sample between two deposits of matrix with a drying step in between. Such "spot-and-dry" approaches would typically allow little time for the matrix molecules to react with proteins. Third, it is assumed that the proteins identified in this study have cysteine residues that are present in their oxidized form, i.e., disulfide bonds. Presumably, SA acts to reduce such disulfide bonds, however, such reduction would depend on the accessibility of the disulfide bonds, which can vary across different proteins. In addition, disulfide bond accessibility may also affect the degree to which a particular protein unfolds or denatures under protein extraction conditions. The potential variability caused by sample preparation (including protein extraction), matrix selection, and even spotting technique may result in significant variations in chemical reactivity of matrices and its impact on the MS spectra obtained.

Fragmentation Efficiency of the Polypeptide Backbone and Its Impact on Top-Down Analysis

In the current study, we observed significant variation in the fragmentation efficiency of the polypeptide chain of protein ions. Among those proteins that fragmented efficiently, we observed fragmentation at multiple cleavage sites. Fragment ions generated from multiple cleavages appeared at the lower m/z range and the cleavages occurred adjacent to D, E, or P residues. We reported similar observations in the fragmentation of intact protein ions of *Campylobacter* [46, 56, 57]. The higher than normal laser fluence used in these experiments (also noted by other researchers [45]) undoubtedly contributed to the increased fragmentation efficiency of the protein ions. The choice of MALDI matrix may also play a significant role. HCCA is considered a "hot" matrix, i.e., generating multiply charged protein ions at a higher intensity and may also facilitate fragmentation of protein ions by increased energy deposition. SA is considered to be a "cold" matrix, i.e., generating primarily singly charged ions and less fragmentation. However, in the current study, we observed
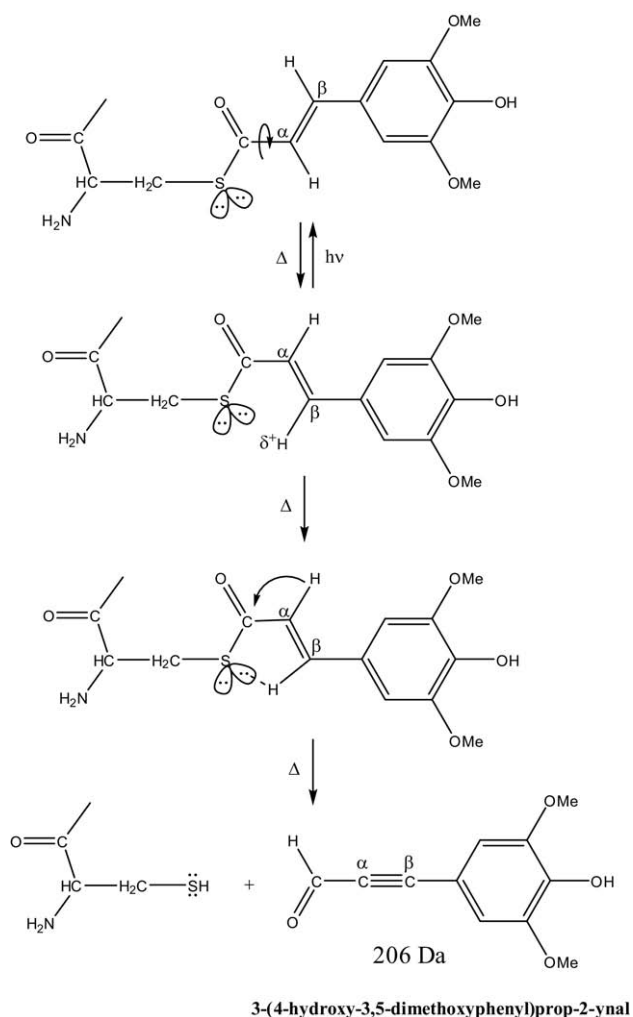
that combining high laser fluence with SA as the matrix resulted in significant protein ion fragmentation. The strong absorption of SA at the laser wavelength of 355 nm may account for its utility for protein ion fragmentation in these studies.

In the current study, protein biomarkers were identified using web-based software, developed in-house, that compares the m/z of MS/MS fragment ions to the m/z of in silico fragment ions on the assumption that the polypeptide backbone fragments only once. The possibility of multiple cleavages of the polypeptide backbone was not incorporated into the in silico fragment ion database. As a consequence, our software in its current form cannot correctly identify fragment ions generated from multiple cleavages of the polypeptide backbone, should they exist. Such fragment ions must be identified from manual inspection after the software has identified the highest scoring protein sequence. Further refinements of our software may include incorporation of multi-cleavage in silico fragment ions into the database [46, 49].

Dissociative Loss of Thiol Ester-Linked SA

As noted previously, it is believed that thiol ester-linked SA may undergo *cis-trans* isomerization of the *trans* vinyl bond as part of the photocycle in the solution phase [53]. SA exhibits a strong absorption in the near-UV with absorption maxima at ~365 nm in thin films [63] and at ~320 nm in solution [64]. In the current study, the laser desorption/ionization wavelength was 355 nm. Gas-phase absorption studies of SA have not been reported, however, given the high laser fluence used in these experiments to facilitate protein fragmentation, it is not unreasonable to assume gas-phase photon absorption may occur by the covalently bound SA [50–54]. Presumably, if photon absorption does occur, it may lead to isomerization of the *trans* vinyl bond to the *cis* configuration as suggested in Scheme 4. Such a photo-induced isomerization would occur during the 5 ns desorption/ionization laser pulse.

The observed dissociative loss of 206 Da from a precursor ion would occur post-source, i.e., after the laser vaporization/desorption pulse and after the pulsed acceleration of ions from the source. Post-source dissociations are statistical dissociations, i.e., internal energy is redistributed amongst all the ro-vibrational degrees-of-freedom of the molecular ion leading to fragmentation at the lowest energy dissociation channels. Statistical dissociations often result in hydrogen migrations and/or rearrangements of the molecular ion before fragmentation as this leads to the lowest energy pathway to dissociation. Following this principle, dissociative loss of the thiol ester-linked SA is likely preceded by a concerted, multi-atom rearrangement from the *trans* configuration of the attached SA as outlined in Scheme 4. We propose a five-membered transition-state involving the sulfur and carbonyl carbon of the thiol ester bond, the α-carbon, the β-carbon

3-(4-hydroxy-3,5-dimethoxyphenyl)prop-2-ynal

Scheme 4

and β-hydrogen. From this pentacyclic transition-state, proton abstraction of the labile β-hydrogen by one of the lone pairs of the sulfur atom leads to formation of a triple bond between the α- and β-carbons, migration of the α-hydrogen to the carbonyl carbon to form an aldehyde and cleavage of the thiol ester bond. The transition-state and rearrangement would be facilitated by the internal energy of the molecular ion, the gas-phase basicity of the sulfur lone pairs and the gas-phase acidity of the β-hydrogen. The stable leaving group of this rearrangement/dissociation would be 3-(4-hydroxy-3,5-dimethoxyphenyl)prop-2-ynal, whose mass of 206 Da is consistent with the mass of the observed dissociative loss.

Conclusions

We report the covalent attachment via a thiol ester bond of sinapinic acid (SA) to the cysteine residues of proteins from bacterial cell lysates prepared for analysis by MALDI. Three cysteine/disulfide bond-containing proteins, HdeA, HdeB, and homeobox (or YbgS) from a

pathogenic and a nonpathogenic strain of *E. coli* were identified by MS, MS/MS and top-down proteomics as having covalently attached SA. SA did not appear to bind to non-cysteine-containing proteins. HCCA did not appear to covalently bind to either cysteine-containing proteins or non-cysteine-containing proteins. The differential reactivity of SA and HCCA toward cysteine residues may be utilized in the identification of unknown cysteine-containing proteins. Gas-phase dissociative loss of the thiol ester-linked SA is proposed to occur via formation of a pentacyclic transition-state comprised of the cysteine sulfur and the carbonyl carbon of the thiol ester bond, the α-carbon, the β-carbon, and β-hydrogen. Proton abstraction of the β-hydrogen by one of the lone pairs of the cysteine sulfur would initiate a series of bond-breaking and bond-forming rearrangements resulting in the dissociative loss of 3-(4-hydroxy-3,5-dimethoxyphenyl)prop-2-ynal. Finally, we report evidence of fragment ions generated from multiple cleavages of the polypeptide backbone of intact protein ions.

Acknowledgments

References

1. Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F. Matrix-Assisted Ultraviolet Laser Desorption of Nonvolatile Compounds. *Int. J. Mass Spectrom. Ion Processes* **1987,** *78,* 53–68.
2. Tanaka, K.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T. Detection of High Mass Molecules by Laser Desorption Time-of-Flight Mass Spectrometry. *Proceedings of the 2nd Japan-China Joint Symposium on Mass Spectrometry;* Osaka, Japan, September, 1987; Abstract p. 185–188.
3. Wu, K. J.; Odom, R. W. Characterizing Synthetic Polymers by MALDI MS. *Anal. Chem.* **1998,** *70,* 456A–461A.
4. Zucht, H. D.; Lamerz, J.; Khamenia, V.; Schiller, C.; Appel, A.; Tammen, H.; Crameri, R.; Selle, H. Datamining Methodology for LC-MALDI-MS Based Peptide Profiling. *Comb. Chem. High Throughput Screen.* **2005,** *8,* 717–723.
5. Hardouin, J. Protein Sequence Information by Matrix-Assisted Laser Desorption/Ionization In-Source Decay Mass Spectrometry. *Mass Spectrom. Rev.* **2007,** *26,* 672–682.
6. Fenselau, C.; Demirev, P. A. Characterization of Intact Microorganisms by MALDI Mass Spectrometry. *Mass Spectrom. Rev.* **2001,** *20,* 157–171.
7. Lay, J. O. Jr. MALDI-TOF Mass Spectrometry of Bacteria. *Mass Spectrom. Rev.* **2001,** *20,* 172–194.
8. Chaurand, P.; Norris, J. L.; Cornett, D. S.; Mobley, J. A.; Caprioli, R. M. New Developments in Profiling and Imaging of Proteins from Tissue Sections by MALDI Mass Spectrometry. *J. Proteome Res.* **2006,** *5,* 2889–2900.
9. Meng, Z.; Simmons-Willis, T. A.; Limbach, P. A. The Use of Mass Spectrometry in Genomics. *Biomol. Eng.* **2004,** *21,* 1–13.
10. König, S.; Kollas, O.; Dreisewerd, K. Generation of Highly Charged Peptide and Protein Ions by Atmospheric Pressure Matrix-Assisted Infrared Laser Desorption/Ionization Ion Trap Mass Spectrometry. *Anal. Chem.* **2007,** *79,* 5484–5488.
11. Li, B.; An, H. J.; Hedrick, J. L.; Lebrilla, C. B. Collision-Induced Dissociation Tandem Mass Spectrometry for Structural Elucidation of Glycans. *Methods Mol. Biol.* **2009,** *534,* 133–145.
12. Ens, W.; Standing, K. G. Hybrid Quadrupole/Time-of-Flight Mass Spectrometers for Analysis of Biomolecules. *Methods Enzymol.* **2005,** *402,* 49–78.
13. Chernushevich, I. V.; Loboda, A. V.; Thomson, B. A. An Introduction to Quadrupole-Time-of-Flight Mass Spectrometry. *J. Mass Spectrom.* **2001,** *36,* 849–865.
14. Yoo, C.; Patwa, T. H.; Kreunin, P.; Miller, F. R.; Huber, C. G.; Nesvizhskii, A. I.; Lubman, D. M. Comprehensive Analysis of Proteins of pH Fractionated Samples Using Monolithic LC/MS/MS, Intact MW Measurement, and MALDI-QIT-TOF MS. *J. Mass Spectrom.* **2007,** *42,* 312–334.
15. Albrethsen, J. Reproducibility in Protein Profiling by MALDI-TOF Mass Spectrometry. *Clin. Chem.* **2007,** *53,* 852–858.
16. Medzihradszky, K. F.; Campbell, J. M.; Baldwin, M. A.; Falick, A. M.; Juhasz, P.; Vestal, M. L.; Burlingame, A. L. The Characteristics of Peptide Collision-Induced Dissociation Using a High-Performance MALDI-TOF/TOF Tandem Mass Spectrometer. *Anal. Chem.* **2000,** *72,* 552–558.
17. Suckau, D.; Resemann, A.; Schuerenberg, M.; Hufnagel, P.; Franzen, J.; Holle, A. A Novel MALDI LIFT-TOF/TOF Mass Spectrometer for Proteomics. *Anal. Bioanal. Chem.* **2003,** *376,* 952–965.
18. Demirev, P. A.; Fenselau, C. Mass Spectrometry in Biodefense. *J. Mass Spectrom.* **2008,** *43*(11), 1441–1457.
19. Claydon, M. A.; Davey, S. N.; Edwards-Jones, V.; Gordon, D. B. The Rapid Identification of Intact Microorganisms Using Mass Spectrometry. *Nat. Biotechnol.* **1996,** *14*(11), 1584–1586.
20. Cain, T. C.; Lubman, D. M.; Weber, W. J. Jr. Differentiation of Bacteria Using Protein Profiles from Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1994,** *8,* 1026–1030.
21. Krishnamurthy, T.; Ross, P. L.; Rajamani, U. Detection of Pathogenic and Nonpathogenic Bacteria by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1996,** *10,* 883–888.
22. Krishnamurthy, T.; Ross, P. L. Rapid Identification of Bacteria by Direct Matrix-Assisted Laser Desorption/Ionization Mass Spectrometric Analysis of Whole Cells. *Rapid Commun. Mass Spectrom.* **1996,** *10,* 1992–1996.
23. Holland, R. D.; Wilkes, J. G.; Rafii, F.; Sutherland, J. B.; Persons, C. C.; Voorhees, K. J.; Lay, J. O. Jr. Rapid Identification of Intact Whole Bacteria Based on Spectral Patterns Using Matrix-Assisted Laser Desorption/Ionization with Time-of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1996,** *10,* 1227–1232.
24. Holland, R. D.; Duffy, C. R.; Rafii, F.; Sutherland, J. B.; Heinze, T. M.; Holder, C. L., Voorhees, K. J.; Lay, J. O. Jr. Identification of Bacterial Proteins Observed in MALDI TOF Mass Spectra from Whole Cells. *Anal Chem.* **1999,** *71,* 3226–3230.
25. Arnold, R.; Reilly, J. Fingerprint Matching of *E. coli* Strains with Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry of Whole Cells Using a Modified Correlation Approach. *Rapid Commun. Mass Spectrom.* **1998,** *12,* 630–636.
26. Welham, K.; Domin, M.; Scannell, D.; Cohen, E.; Ashton, D. The Characterization of Microorganisms by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1998,** *12,* 176–180.
27. Haag, A.; Taylor, S.; Johnston, K.; Cole, R. Rapid Identification and Speciation of *Haemophilus* Bacteria by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *J. Mass Spectrom.* **1998,** *33,* 750–756.
28. Wang, Z.; Russon, L.; Li, L.; Roser, D.; Long, S. R. Investigation of Spectral Reproducibility in Direct Analysis of Bacteria Proteins by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1998,** *12,* 456–464.
29. Dai, Y.; Li, L.; Roser, D.; Long, S. R. Detection and Identification of Low-Mass Peptides and Proteins from Solvent Suspensions of *Escherichia coli* by High Performance Liquid Chromatography Fractionation and Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1999,** *13,* 73–78.
30. Ramirez, J.; Fenselau, C. Factors Contributing to Peak Broadening and Mass Accuracy in the Characterization of Intact Spores Using Matrix-Assisted Laser Desorption/Ionization Coupled with Time-of-Flight Mass Spectrometry. *J. Mass Spectrom.* **2001,** *36,* 929–936.
31. Whiteaker, J.; Karns, J.; Fenselau, C.; Perdue, M. L. Analysis of *Bacillus anthracis* Spores in Milk Using Mass Spectrometry. *Anal. Chem.* **2004,** *1,* 185–194.
32. Mandrell, R. E.; Harden, L. A.; Bates, A. H.; Miller, W. G.; Haddon, W. F.; Fagerquist, C. K. Speciation of *Campylobacter coli, C. jejuni, C. helveticus, C. lari, C. sputorum,* and *C. upsaliensis* by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Applied Environ. Microbiol.* **2005,** *71,* 6292–6307.
33. Fagerquist, C. K.; Miller, W. G.; Harden, L. A.; Bates, A. H.; Vensel, W. H.; Wang, G.; Mandrell, R. E. Genomic and Proteomic Identification of a DNA-Binding Protein Used in the "Fingerprinting" of *Campylobacter* Species and Strains by MALDI-TOF-MS Protein Biomarker Analysis. *Anal. Chem.* **2005,** *77,* 4897–4907.
34. Fagerquist, C. K.; Bates, A. H.; Heath, S.; King, B. C.; Garbus, B. R.; Harden, L. A.; Miller, W. G. Subspeciating *Campylobacter jejuni* by Proteomic Analysis of Its Protein Biomarkers and Their Post-Translational Modifications. *J. Proteome Res.* **2006,** *5,* 2527–2538.
35. Jarmon, K. H.; Cebula, S. T.; Sinapinic acidenz, A. J.; Petersen, C. E.; Valentine, N. B.; Kingsley, M. T.; Wahl, K. L. An Algorithm for Automated Bacterial Identification Using Matrix-Assisted Laser De-

sorption/Ionization Mass Spectrometry. *Anal. Chem.* **2000,** *72,* 1217–1223.

36. Wahl, K. L.; Wunschel, S. C.; Jarman, K. H.; Valentine, N. B.; Petersen, C. E.; Kingsley, M. T.; Zartolas, K. A.; Saenz, A. J. Analysis of Microbial Mixtures by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *Anal. Chem.* **2002,** *74,* 6191–6199.

37. Wunschel, S. C.; Jarman, K. H.; Petersen, C. E.; Valentine, N. B.; Wahl, K. L.; Schauki, D.; Jackman, J.; Nelson, C. P.; White, E., 5th. Bacterial Analysis by MALDI-TOF Mass Spectrometry: An Inter-Laboratory Comparison. *J. Am. Soc. Mass Spectrom.* **2005,** *16,* 456–462.

38. Demirev, P. A.; Ho, Y.-P., Ryzhov, V.; Fenselau, C. Microorganism Identification by Mass Spectrometry and Protein Database Searches. *Anal. Chem.* **1999,** *71,* 2732–2738.

39. Peneda, F. J.; Lin, J. S.; Fenselau, C.; Demirev, P. A. Testing the Significance of Microorganism Identification by Mass Spectrometry and Proteome Database Search. *Anal. Chem.* **2000,** *72,* 3739–3744.

40. Demirev, P. A.; Lin, J. S.; Peneda, F. J.; Fenselau, C. Bioinformatics and Mass Spectrometry for Microorganism Identification: Proteome-Wide Post-Translational Modifications and Database Search Algorithms for Characterization of Intact *H. pylori. Anal. Chem.* **2001,** *73,* 4566–4573.

41. Yao, Z.-P.; Demirev, P. A.; Fenselau, C. Mass Spectrometry-Based Proteolytic Mapping for Rapid Virus Identification. *Anal. Chem.* **2002,** *74,* 2529–2534.

42. Peneda, F. J.; Antoine, M. D.; Demirev, P. A.; Feldman, A. B.; Jackman, J.; Longenecker, M., Lin, J. S. Microorganism Identification by Matrix-Assisted Laser/Desorption Ionization Mass Spectrometry and Model-Derived Ribosomal Protein Biomarkers. *Anal. Chem.* **2003,** *75,* 3817–3822.

43. Demirev, P. A.; Ramirez, J.; Fenselau, C. Tandem Mass Spectrometry of Intact Proteins for Characterization of Biomarkers from *Bacillus cereus* T spores. *Anal. Chem.* **2001,** *73,* 5725–5731.

44. Lin, M.; Campbell, J. M.; Mueller, D. R.; Wirth, U. Intact Protein Analysis by Matrix-Assisted Laser Desorption/Ionization Tandem Time-of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2003,** *17,* 1809–1814.

45. Demirev, P. A.; Feldman, A. B.; Kowalski, P.; Lin, J. S. Top-Down Proteomics for Rapid Identification of Intact Microorganisms. *Anal. Chem.* **2005,** *77,* 7455–7461.

46. Fagerquist, C. K.; Garbus, B. R.; Williams, K. E.; Bates, A. H.; Boyle, S.; Harden, L. A. Web-Based Software for Rapid "Top-Down" Proteomic Identification of Protein Biomarkers with Implications for Bacterial Identification. *Appl. Environ. Microbiol.* **2009,** *75,* 4341–4353.

47. Fagerquist, C. K.; Garbus, B. R.; Williams, K. E.; Bates, A. H.; Boyle, S.; Harden, L. A.; Miller, W. G.; Mandrell, R. E. Rapid Identification of *E. coli* O157:H7 by "Top-Down" Proteomics Using MALDI-TOF/TOF Mass Spectrometry. Oral Presentation Given on June 2nd 2009 at the 57th American Society of Mass Spectrometry Conference (Philadelphia, PA). Extended abstract published September 2009 in the Proceedings of the 57th ASMS Conference (on-DVD: A091307.8125VER. 1.pdf).

48. Perna, N. T.; Plunkett, G. III; Burland, V.; Mau, B.; Glasner, J. D.; Rose, D. J.; Mayhew, G. F.; Evans, P. S.; Gregor, J.; Kirkpatrick, H. A.; Pósfai, G.; Hackett, J.; Klink, S.; Boutin, A.; Shao, Y.; Miller, L.; Grotbeck, E. J.; Davis, N. W.; Lim, A.; Dimalanta, E. T.; Potamousis, K. D.; Apodaca, J.; Anantharaman, T. S.; Lin, J.; Yen, G.; Schwartz, D. C.; Welch, R. A.; Blattner, F. R. Genome Sequence of Enterohemorrhagic *Escherichia coli* O157:H7. *Nature* **2001,** *409*(6819), 529–533.

49. Fagerquist, C. K.; Garbus, B. R.; Miller, W. G.; Williams, K. E.; Bates, A. H.; Yee, E.; Boyle, S.; Harden, L. A.; Cooley, M. B.; Mandrell, R. E. Rapid Identification of Protein Biomarkers of *E. coli* O157:H7 by MALDI-TOF-TOF Mass Spectrometry and Top-Down Proteomics, un-published (submitted).

50. van der Horst, M. A.; Arents, J. C.; Kort, R.; Hellingwerf, K. J. Binding, Tuning, and Mechanical Function of the 4-Hydroxy-Cinnamic Acid

Chromophore in Photoactive Yellow Protein. *Photochem. Photobiol. Sci.* **2007,** *6,* 571–579.

51. Hoff, W. D.; Devreese, B.; Fokkens, R.; Nugteren-Roodzant, I. M.; Van Beeumen, J.; Nibbering, N.; Hellingwerf, K. J. Chemical Reactivity and Spectroscopy of the Thiol Ester-Linked p-Coumaric Acid Chromophore in the Photoactive Yellow Protein from *Ectothiorhodospira halophila.* *Biochemistry* **1996,** *35,* 1274–1281.

52. Baca, M.; Borgstahl, G. E.; Boissinot, M.; Burke, P. M.; Williams, D. R.; Slater, K. A.; Getzoff, E. D. Complete Chemical Structure of Photoactive Yellow Protein: Novel Thioester-Linked 4-Hydroxycinnamyl Chromophore and Photocycle Chemistry. *Biochemistry* **1994,** *33,* 14369–14377.

53. Hoff, W. D.; Düx, P.; Hård, K.; Devreese, B.; Nugteren-Roodzant, I. M.; Crielaard, W.; Boelens, R.; Kaptein, R.; van Beeumen, J.; Hellingwerf, K. J. Thiol Ester-Linked p-Coumaric Acid as a New Photoactive Prosthetic Group in a Protein with Rhodopsin-Like Photochemistry. *Biochemistry* **1994,** *33,* 13959–13962.

54. Van Beeumen, J. J.; Devreese, B. V.; Van Bun, S. M.; Hoff, W. D.; Hellingwerf, K. J.; Meyer, T. E.; McRee, D. E.; Cusanovich, M. A. Primary Structure of a Photoactive Yellow Protein from the Phototrophic Bacterium *Ectothiorhodospira halophila,* with Evidence for the Mass and the Binding Site of the Chromophore. *Protein Sci.* **1993,** *2,* 1114–1125.

55. Taylor, J. A.; Johnson, R. S. Implementation and Uses of Automated De Novo Peptide Sequencing by Tandem Mass Spectrometry. *Anal. Chem.* **2001,** *73,* 2594–2604.

56. Fagerquist, C. K.; Williams, K. E.; Bates, A. H. Identification of Food-borne Bacteria by High Energy Collision-Induced Dissociation of Their Protein Biomarkers by MALDI Tandem-Time-of-Flight Mass Spectrometry. *Proceedings of the 55th ASMS Conference on Mass Spectrometry and Allied Topics*; Indianapolis, IN, June, 2007.

57. C. K. Fagerquist. Identification of Foodborne Bacteria by MALDI-TOF-TOF Analysis of Protein Biomarkers. *Proceedings of the 121st AOAC Annual Meeting*; Anaheim, CA, 2007.

58. Mandrell, R. E.; Harden, L. A.; Horn, S. T.; Haddon, W. F.; Miller, W. G. Analysis of *E. coli* Environmental and Diarrheal Isolates by MALDI-TOF Mass Spectrometry: Identification of Potential Biomarker Ions and a Mutation in a Gene Encoding a Biomarker Ion. *Proceedings of the American Society of Microbiology*; Los Angeles, CA, May, 2000; Poster C-177.

59. Mazzeo, M. F.; Sorrentino, A.; Gaita, M.; Cacace, G.; Di Stasio, M.; Facchiano, A.; Comi, G.; Malorni, A.; Siciliano, R. A. Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry for the Discrimination of Food-Borne Microorganisms. *Appl. Environ. Microbiol.* **2006,** *72*(2), 1180–1189.

60. Yang, H.; Liu, N.; Qiu, X.; Liu, S. A New Method for Analysis of Disulfide-Containing Proteins by Matrix-Assisted Laser Desorption Ionization (MALDI) Mass Spectrometry. *J Am. Soc. Mass Spectrom.* **2009,** *20,* 2284–2293.

61. Happersberger, H. P.; Bantscheff, M.; Barbirz, S.; Glocker, M. O. Multiple and Subsequent MALDI-MS On-Target Chemical Reactions for the Characterization of Disulfide Bonds and Primary Structures of Proteins. *Methods Mol. Biol.* **2000,** *146,* 167–184.

62. Spiess, C.; Happersberger, H. P.; Glocker, M. O.; Spiess, E.; Rippe, K.; Ehrmann, M. Biochemical Characterization and Mass Spectrometric Disulfide Bond Mapping of Periplasmic α-Amylase MalS of *Escherichia coli. J. Biol Chem.* **1997,** *272,* 22125–22133.

63. Allwood, D. A.; Dreyfus, R. W.; Perera, I. K.; Dyer, P. E. UV Optical Absorption of Matrices Used for Matrix-Assisted Laser Desorption/ Ionization. *Rapid Commun. Mass Spectrom.* **1996,** *10,* 1575–1578.

64. Low, W.; Kang, J.; DiGruccio, M.; Kirby, D.; Perrin, M.; Fischer, W. H. MALDI-MS Analysis of Peptides Modified with Photolabile Aryl-Azido Groups. *J. Am. Soc. Mass Spectrom.* **2004,** *15,* 1156–1160.