Complementary b/y fragment ion pairs from post-source decay of metastable YahO for calibration of MALDI-TOF-TOF-MS/MS

Clifton K. Fagerquist* and William J. Zaragoza

Produce Safety & Microbiology Research Unit, Western Regional Research Center, Agricultural Research Service, U.S. Department of Agriculture, 800 Buchanan Street, Albany, California 94710

Running title: Complementary b/y fragment ion pairs of metastable YahO

*Address correspondence to Clifton K. Fagerquist, clifton.fagerquist@ars.usda.gov.

Tel. 510-559-5691, FAX: 510-559-6162.

Abstract

Complementary b/y fragment ion pairs from post-source decay (PSD) of metastable YahO protein ion were evaluated for use in the calibration of MALDI-TOF-TOF for tandem mass spectrometry (MS/MS). The *yahO* gene from pathogenic *Escherichia coli* O157:H7 strain EDL933 was cloned into a pBAD18 plasmid vector and transferred to a *yahO* gene knockout (Δ *yahO*) of non-pathogenic *Escherichia coli* strain K-12. YahO protein was over-expressed by overnight culturing of the K-12 mutant on Luria-Bertani agar (LBA) supplemented with arabinose. MS/MS-PSD of the YahO protein ion results in polypeptide backbone cleavage on the C-terminal side of seven aspartic acid (D) residues of the mature protein sequence resulting in a series of prominent b/y fragment ion pairs. Interestingly, we observe an unusual oscillatory pattern in the fragmentation efficiency of cleavage sites across the linear polypeptide chain of YahO. The instrument was calibrated for MS/MS-PSD using at least six of twelve of the most intense b- and y-type fragment ions of YahO. Calibration was tested on the disulfide bond-reduced B-subunit of Shiga toxin 2 from *Escherichia coli* O157:H7 strain EDL933 cultured overnight on LBA supplemented with ciprofloxacin. Two non-simultaneous calibrant/analyte acquisition scenarios were evaluated: 1. YahO and B-subunit on separate (but adjacent) spot locations; 2. YahO and the B-subunit co-located on the same spot. We observed no significant difference in mass accuracy for co-location vs. adjacent spot analysis in spite of a significant drop in ion intensity of both calibrant and analyte for co-location analysis.

Introduction

Matrix-assisted laser desorption/ionization time-of-flight-time-of-flight (MALDI-TOF-TOF) platforms are increasingly utilized for tandem mass spectrometry (MS/MS) of intact proteins. Two strategies have been utilized. The first strategy utilizes in source-decay (ISD) of a protein followed by MS/MS analysis of the resulting fragment ions for detection and identification of 2nd generation fragment ions. This has been described as T3 sequencing [1]. Sequence coverage occurs primarily near the N- and C-termini. Like Edman sequencing, the limitation of this approach is that it can only be used to analyze *pure* proteins. Analysis of protein mixture (e.g. bacterial cell lysate) is not possible because fragment ions generated by ISD cannot be attributed unambiguously to a specific protein precursor ion.

The second strategy involves TOF separation/isolation of protein ions based on their arrival time at a mass gate followed by fragmentation, by either collision-induced dissociation (CID) or post-source decay (PSD), and analysis of the resulting fragment ions [2]. Fragment ions are often the result of polypeptide backbone cleavage on the C-terminal side of aspartic acid (D) and glutamic acid (E) residues and the N-terminal side of proline (P) residues. The advantage of this approach is that it is possible to analyze proteins from a mixture (e.g. bacterial cell lysate) as each protein precursor ion is isolated prior to its fragmentation, and the resulting fragment ions can be attributed to a specific precursor ion as with any classical MS/MS experiment.

Instrument calibration is obviously critical to any MS or MS/MS experiment. As new applications are developed for mass spectrometry platforms, calibrants are needed to optimize instrument mass accuracy. Gruendling et al. reported using polyalanine as an MS and MS/MS calibrant for MALDI-TOF-TOF in both positive and negative ion mode. The reported mass

range in MS mode was $\sim m/z$ 700 to 3500. In MS/MS mode with a precursor ion of m/z 2432, c- and y-type fragment ions were observed in the mass range of $\sim m/z$ 100 to 2300 [3].

Calibrants for MS/MS-PSD analysis of protein ions are limited by the fact that their fragmentation efficiency is often dependent on the primary sequence of the protein as well as secondary and tertiary structures. In addition, PSD fragment ions from a metastable protein precursor ion have a higher mass-to-charge (m/z) than fragment ions from a peptide calibrant, and thus may not be fully resolved by the 2nd TOF mass analyzer. In consequence, the average m/z of the fragment ions must be used for instrument calibration. Using a commercially available protein and reagents, we had previously demonstrated the use of alkylated thioredoxin as a calibrant for MS/MS-PSD ($\sim m/z$ 5k-10k) for a MALDI-TOF-TOF platform as an alternative to peptide calibrants that generate fragment ions in the lower m/z range [4].

In the current work, a YahO protein expression plasmid was developed and tested as a MS/MS calibrant. The fragment ions of metastable YahO span a wide m/z range (1500 to 7000). The YahO protein from pathogenic *Escherichia coli* O157:H7 strain EDL933 had been previously shown to fragment efficiently by MS/MS-PSD even when ionized from an unfractionated bacterial cell lysate of the wild type strain [5]. The facile fragmentation efficiency of YahO is due, in part, to seven D residues that are fortuitously placed in the primary sequence resulting in a series of complementary b/y fragment ion pairs. The combined intensity of these complementary fragment ion pairs reveal an unusual oscillatory pattern in fragmentation efficiency across the polypeptide chain. The YahO calibrant was tested for MS/MS-PSD analysis of the B-subunit of Shiga toxin 2 from *Escherichia coli* O157:H7 and found to provide excellent calibration characteristics for both co-location and adjacent spot analysis. Portions of this work were presented at the 64th Conference on Mass Spectrometry and Allied Topics [6].

Material and Methods

Bacterial strains, plasmids, and growth media

E. coli strains and mutants used in this study are listed in **Table 1** [7-10]. For routine culturing, strains were grown in Luria-Bertani broth (LBB) at 27 °C or 37 °C statically or with shaking at 200 rpm, depending on the application. Antibiotic concentrations were used as follows: ampicillin (Amp) 100 µg mL⁻¹, carbenicillin (Cb) 50 µg mL⁻¹, and kanomycin (Km) 50 µg mL⁻¹.

Strain and gene selection

Strains are listed in **Table 1**. Strain MG1655 K-12 was selected as the host train. The *yahO* nucleotide sequence in *E. coli* strain EDL-933 differs from the homologue in MG1655 K-12 by six nucleotides and the resulting proteins differ by only two amino acids (14F ↔ 14L and 86D ↔ 86N, respectively). However, the substitution at residue 14 is part of the signal peptide which is not present in the mature protein either from the wild type or from the plasmid. The YahO protein is an uncharacterized protein with a signal peptide that suggests that it is destined for sub-cellular location in the periplasm. The YahO protein from EDL-933 yields a superior fragmentation pattern by post-source decay (PSD) with a broader spectrum of peaks that can be used for calibration by MS/MS. To remove the potential for interference from the expression of the endogenous K-12 *yahO* gene product, this gene was deleted from the K-12 host using the Lambda Red Recombination system as previously described [9].

Polymerase chain reaction (PCR) amplification

Primers used in this study are listed in **Table 2**. Linear or plasmid DNA was purified with either the Illustra GFX PCR DNA and gel band purification kit (*GE Healthcare and Life Sciences*, catalogue number 28-9034-70), the QIAamp DNA mini kit (Qiagen, catalogue # 51304), or the QIAGEN Plasmid Mini Kit (Qiagen, catalogue # 12123) as appropriate, according to manufacturer specifications. Amplification of the Km cassette with appended “arms” homologous to *yahO* from *E. coli* MG1655 K-12 was carried out with the pKD4 plasmid template using PFX Platinum Polymerase (Life Technologies, Catalogue number 11708013) following manufacturer specifications for reagent ratios. PCR was performed on a Tetrad 2 (Bio-Rad, Hercules, CA). Confirmatory PCRs were carried out with Taq DNA polymerase (New England BioLabs, catalogue # M0273L) and PFX polymerase, respectively. Cloning of the *yahO* gene from strain EDL-933 was carried out with PFX polymerase using the thermocycler conditions used for verification.

K-12 mutant

All transformations were carried out via electroporation. Briefly, cells were washed three times in ice-cold 10% glycerol, pelleted, and resuspended in molecular biology grade water with linearized DNA, pulsed, and plated on solid agar plates with the appropriate antibiotics. Plates were then incubated at temperatures appropriate for each step (27 °C or 37 °C). The Km cassette was amplified from the pKD4 plasmid, generating a linear piece of DNA that contained the Km cassette flanked by regions of homology to the *E. coli* K-12 MG1655 *yahO* gene. Deletion of

the gene encoding YahO was achieved by replacing the open rereading encoding *yahO* with the kanomycin. DNA was transformed into the K-12 strain bearing the helper plasmid pKD46. Transformants were selected on LB agar plates containing kanomycin at 50 $\mu\text{g mL}^{-1}$. Insertion was assayed with confirmatory primers listed in Table 2. The kanomycin marker was removed using FLP-mediated recombinase in plasmid pCP20 [9].

pBAD *yahO* expression construct

Plasmid pBAD18-Cb (generously provided by Dr. Michelle Q. Carter) was purified and double-digested with HindIII and KpnI at 37 °C for 1 hour (New England BioLabs, catalogue # R0104 and R0142, respectively). This digest was then quenched with Proteinase K and purified. The *yahO* gene was cloned from *E. coli* strain EDL933 using the appropriate primers from **Table 2**. The PCR product is then ligated into the linearized pBAD18-Cb plasmid using T4 DNA ligase (New England BioLabs, catalogue # M0202S) per manufacturer's recommendations, resulting in plasmid pWZ19. pWZ19 is then transformed into WJZ20, resulting in a K-12 $\Delta yahO$ construct that is missing its endogenous *yahO* gene, but has been complemented with *yahO* from strain EDL-933, borne by the pBAD18-Cb vector. These transformants are plated on Cb containing media and confirmed via PCR and MS.

Mass spectrometry

Strains WJZ20 and RM1272 were grown statically over night at 37 °C on LB agar supplemented with carbinecillin at 50 $\mu\text{g mL}^{-1}$ and 0.3% L-arabinose (for WJZ20), or

ciprofloxacin at 20 ng mL⁻¹ (for RM1272). A flush, 1 µL loop of cells were harvested for both strains, placed in an O-ring sealed screw-cap tube with 300 µL of HPLC-grade H₂O (Fisher Scientific) and ~40 mg of 0.1 mm zirconia/silica beads, and mechanically lysed in a bead-beater (Biospec Products Inc., Bartlesville, OK, USA) for one minute. Cells were then centrifuged for 2 min at 13000 rpm. Supernatant was used for subsequent analysis by mass spectrometry analysis.

Mass spectrometry (MS) and tandem mass spectrometry (MS/MS) data was collected using a 4800 MALDI-TOF-TOF instrument (Sciex) as previously described [5, 11]. Briefly, 0.6 µL aliquots of bacterial cell lysate supernatant were spotted and allowed to dry on a 384-spot stainless steel target. Dried sample spots were then overlaid with a 0.6 µL aliquot of saturated solution of sinapinic acid dissolved in 67% water, 33% acetonitrile and 0.2 % trifluoroacetic acid which was also allowed to dry.

MS linear-mode external calibration was performed using the +1 and +2 charge states of cytochrome-C, lysozyme and myoglobin (Sigma-Aldrich, St. Louis, MO). External calibration of the MS/MS reflectron-mode was performed using the supernatant of the unfractionated bacterial cell lysate of K-12 mutant containing the *yahO* gene from EDL933. Fractionation was not necessary as the only protein detected was the highly abundant YahO (+1 and +2) and the matrix adducted YahO. Calibration was performed using at least six of twelve b- and y-type fragment ions (**Table 3**) resulting from cleavage of the polypeptide backbone on the C-terminal side of seven aspartic acid (D) residues by post-source decay (PSD). For MS/MS, ions were accelerated from the 1st source at 8 kV, mass selected by the timed-ion selector (TIS) based upon their arrival time, decelerated in the collision cell to 1 kV and the resulting fragment ions are re-accelerated in the 2nd source at 15 kV. As no collision gas was introduced into the collision cell,

fragmentation occurs by PSD. After MS/MS calibration with YahO, MS/MS data were collected on the B-subunit of Stx2 at the same spot location as the calibrant or at a separate but adjacent spot location. Average data and associated error (standard deviation) represent acquisition of three technical replicates.

Results and Discussion

Figure 1A shows the MS data of the bacterial cell lysate of pBAD18_*yahO*_EDL933_K-12_*ΔyahO* grown overnight on LBA supplemented with arabinose. Only YahO protein ions appear in the spectrum: $[M+H]^+$ at m/z 7711 (theoretical $[M+H]^+$ m/z 7708.631) and $[M+H]^{+2}$ at m/z 3852. No bacterial proteins are detected. We have observed similar behavior with the pBAD expression plasmid [12]. We ascribe this phenomenon to down-regulation of host proteins as well as preferential ionization of the highly abundant plasmid protein by MALDI to the detriment of less abundant proteins. The peak at m/z 7921 is the sinapinic acid matrix adduct of the YahO protein ion [12-13]. **Figure 1B** shows an expanded m/z range of **Figure 1A**. Shoulder peaks at m/z 7726 and 7742 appear to be the result of oxidation of the two methionine residues at positions 4 and 28 in the mature sequence. Similarly, the matrix adducted peaks at m/z 7921 and 8129 also have shoulder peaks indicating methionine oxidation. **Figure 1C** is the MS/MS-PSD of the YahO protein ion at m/z 7711 (TIS: ± 100 Da). Fragment ions are identified by their m/z , ion type and number. The mature protein sequence (70 residues long with the signal peptide removed) of YahO is shown between **Figures 1B** and **1C**. An asterisk indicates the site of cleavage and the resulting fragment ions are identified above (y-type) and below (b-type) the sequence. Methionine residues are underlined. The mature protein sequence of YahO

has seven aspartic acid (D) residues. PSD is a relatively low energy dissociation technique [14]. In consequence, PSD of a protein precursor ion results in polypeptide backbone fragmentation via the lowest energy fragmentation channel, i.e. C-terminal side of D-residues. Although other polypeptide backbone fragment ions are present, the most abundant fragment ions are the result of this facile dissociation channel.

Cleavage of the polypeptide backbone of metastable YahO protein ion occurs efficiently on the C-terminal side of D-residues and somewhat less efficiently on the C-terminal side of E-residues. The facile fragmentation by D-residues is likely due to its short side-chain that transfers a proton to the polypeptide backbone. A number of reaction mechanisms have been postulated for the “aspartic acid effect” in peptide and protein ions [15-16]. Yu et al. postulated a two-step mechanism involving proton transfer from the D-residue side-chain to the amide nitrogen on its C-terminal side with subsequent attack of the carboxylate anion at the amide carbon resulting in formation of a cyclic anhydride and cleavage of the polypeptide backbone [15]. Paiz et al. postulated a four-center-one-step mechanism where proton transfer and carboxylate anion attack occur simultaneously [16]. Our data would tend to support the mechanism as postulated by Wu et al and/or Paiz et al to account for the “aspartic acid effect” for the singly protonated metastable YahO [15-16].

Figure 2 is a plot of the ratio of the intensities of the fragment ion pair $y_{(70-x)}/b_x$ of YahO, i.e. the two complementary fragment ions (including neutral losses) generated from cleavage of the polypeptide backbone at the same site. Also plotted is the ratio of protonation sites for each $y_{(70-x)}/b_x$ fragment ion pair (opposing Y-axis). The relative intensity of the $y_{(70-x)}/b_x$ fragment ion pair corresponds, in a qualitative way, with the number of protonation sites in each fragment ion. If one assumes that the probability of protonation is roughly equivalent for all thirteen basic sites

(the N-terminus, eleven lysine residues and one histidine residue), the greater the number of protonation sites in a protein fragment, the greater the probability that the fragment will sequester the ionizing proton. With respect to the design of protein calibrants for MALDI-MS/MS-PSD, it is not just the number and location of D-residues that are critical but also the number and location of basic residues that act to sequester the single ionizing proton.

Figure 3 shows a plot of the normalized fragment ion pair intensities (including neutral losses) as a function of polypeptide backbone cleavage site revealing the fragmentation efficiency across the YahO chain. The vertical dotted line denotes the center of the 70-residue polypeptide chain. A general trend is observed where the middle of the chain appears more likely to cleave than near the ends. Overlaying this general trend is an oscillating pattern in the fragmentation efficiency of cleavage sites on the C-terminal side of D-residues. Interestingly, the cleavage site closest to the center of the chain ($y_{34}+b_{36}$) shows a significant decline in efficiency compared to the flanking sites at $y_{38}+b_{32}$ and $y_{28}+b_{42}$. Although the adjacent glutamic acid residue fragmentation channel may reduce the intensity of the $y_{34}+b_{36}$ fragment ion pair, the same could be argued for the glutamic acid residue fragmentation channel adjacent to the $y_{28}+b_{42}$ fragment ion pair. The cleavage sites $y_{49}+b_{21}$ and $y_{23}+b_{47}$, whose positions on the chain are approximately equidistant from the center, have nearly equivalent cleavage efficiency. To the best of our knowledge, this oscillatory behavior has not been reported before for peptides and proteins. It is unclear whether the observed oscillation is a result of vibrational, rotational or larger amplitude motion of the polypeptide chain. Although there have been numerous studies of dissociation of metastable peptide ions using statistical and non-statistical approaches [17-18], there have been little (if any) investigations on the dissociation of low charge state metastable protein ions perhaps due to the computational complexity of such a large molecule. The

practical importance of the observed oscillation is that it directly effects the relative abundance of calibrant fragment ions.

MS/MS calibration using b/y fragment ion pairs of metastable YahO

Figure 4A shows the MS data of the bacterial cell lysate of *E. coli* O157:H7 strain EDL933 grown overnight on LBA supplemented with ciprofloxacin. The dominating ion is the disulfide-reduced Stx2a B-subunit at m/z 7819 ($[M+H]^+$ theoretical m/z 7818.647). **Figure 4B** shows an expanded m/z range of **Figure 4A**. The peak at m/z 8027 is the matrix adducted B-subunit. **Figure 4C** is the MS/MS-PSD of the disulfide-reduced B-subunit at m/z 7819 (TIS: $\pm$ 100 Da). Fragment ions are identified by their m/z , ion type and number. The mature protein sequence (signal peptide removed) of B-subunit is shown between **Figures 4B** and **4C**. An asterisk indicates the cleavage site, and the resulting fragment ions are identified above (y-type) and below (b-type) the sequence. Two cysteine residues and a methionine residue are underlined. The three most prominent fragment ions are the result of polypeptide backbone fragmentation on the C-terminal side of three D-residues. Although the Stx2a B-subunit has five D-residues, one is the penultimate residue at the N-terminus and the other is the terminal residue at the C-terminus. Three less prominent fragment ions are the result of polypeptide backbone cleavage on the C-terminal side of three glutamic acid (E) residues.

Figure 5A shows the MS data of the co-spotted bacterial cell lysates of *E. coli* O157:H7 strain EDL933 grown overnight on LBA supplemented with ciprofloxacin and pBAD18_*yahO*_EDL933_K-12_ Δ *yahO* grown overnight on LBA supplemented with arabinose. The expanded m/z insert clearly shows the YahO and B-subunit ions. We observe a significant

decline in signal intensity for both the calibrant and analyte ion for same spot analysis compared with adjacent spot analysis (**Figures 1A&B** and **4A&B**). As the same amount of calibrant and analyte was added for same spot and adjacent spot analysis, the reason for this concomitant decline in intensity may be due to mutual ionization suppression. Although this phenomenon is well understood in ESI-MS where analytes compete for the finite number of available charges in a charged droplet, it is not as well understood in MALDI. **Figure 5B** shows MS/MS-PSD of the YahO protein ion using a TIS window of $-100\text{ Da}/+50\text{ Da}$. Fragment ions are identified by their m/z , ion type and number. The mature protein sequence (signal peptide removed) of YahO is shown between **Figures 5A** and **5B**. **Figure 5C** shows MS/MS-PSD of the B-subunit protein ion using a TIS window: $-40\text{ Da}/+75\text{ Da}$. As can be seen from **Figures 5B** and **5C** there is little evidence for ion spillover between these two non-simultaneous MS/MS data acquisitions from the same sample spot which highlights the effectiveness of the TIS in isolating the calibrant and analyte precursor ions from one another during MS/MS. MS/MS spectra of this co-spotted analysis is qualitatively similar to adjacent spot MS/MS analysis although adjacent spot analysis has improved signal-to-noise due to a more abundant precursor ion for both analyte and calibrant.

Table 4 shows the average m/z of the three most abundant fragment ions (y_{46} , y_{53} , y_{54}) of the Stx2a B-subunit as analyzed by adjacent spot vs. same spot analysis. The experiment was repeated on two separate days. There is no significant difference in the fragment ion error for adjacent spot vs. same spot analysis. However, there is a significant difference in the fragment ion intensity between same spot and adjacent spot analysis on both days, which is also reflected in the MS data, and may be due to either mutual ionization suppression of YahO and the B-subunit or differences in matrix crystallization (or both). Although the intensity of the fragment

ions of the B-subunit for same spot analysis are $\frac{1}{2}$ to $\frac{1}{3}$ less intense than that of adjacent spot analysis, the fragment ion error does not appear to be affected. Unfortunately, it is not possible with this particular instrument platform to acquire calibrant and analyte data simultaneously with the objective of a post-acquisition data correction although such an approach may be possible on other MALDI-TOF-TOF platforms.

Comparison of the protein calibrant alkylated thioredoxin (AlkTrx) [4] to the YahO calibrant is challenging given the limited number (three) of prominent fragments ions of the B-subunit of Stx2 analyte and the narrow m/z range they fall within (~ 5000 - 6000). However, it would seem that AlkTrx and YahO calibrants appear to be somewhat comparable in terms of performance and both are superior to the peptide calibrant [Glu1]-fibrinopeptide B [4]. The calibrant fragment ions of YahO span a m/z range of 1500 - 7000 whereas the calibrant fragment ions of AlkTrx span a m/z range of ~ 5000 - 9500 . Thus, the YahO calibrant has the advantage of covering a wide m/z range spanning from peptide calibrants to the higher molecular weight AlkTrx calibrant.

Conclusions

Metastable YahO protein ion analyzed by MS/MS-PSD produces complementary b/y fragment ion pairs from polypeptide backbone cleavage on the C-terminal side of seven nearly equally spaced aspartic acid residues in the mature sequence. The relative abundance of the complementary b-type vs. y-type fragment ion scales approximately with the relative number of basic residues present in each complementary fragment ion. The combined intensities of complementary b- and y-type fragment ions shows an unusual oscillatory pattern in

fragmentation efficiency across the linear polypeptide chain. Prominent b- and y-type fragment ions from metastable YahO were used successfully to externally calibrate a MALDI-TOF-TOF for MS/MS-PSD analysis of the B-subunit of Shiga toxin from an unfractionated bacterial cell lysate of *E. coli* O157:H7. There was no detectable difference in mass accuracy between co-location *versus* separate/adjacent spot analysis of calibrant and analyte even though the intensities of calibrant and analyte were significantly reduced for co-located spots. The fragmentation efficiency of singly charged (protonated) metastable YahO as analyzed by MS/MS-PSD presents a rich and (for the most part) uncomplicated series of fragment ions resulting from the fortuitous spacing of aspartic acid and basic residues in the mature protein sequence. As such, it is highly useful both as a mass calibrant and as a model for protein ion fragmentation that highlights the “aspartic acid effect” for polypeptide backbone cleavage of a low charge state protein ion.

Acknowledgements

The YahO mutant *E. coli* strain is available for research to other laboratories after completion of a material transfer agreement. This work was supported by ARS, USDA CRIS Project: 2030-42000-051-00-D.

References

1. Suckau, D., Resemann, A.: T3-sequencing: targeted characterization of the N- and C-termini of undigested proteins by mass spectrometry. *Anal Chem.*, 75, 5817-5824 (2003)
2. 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.* 17, 1809–1814 (2003)
3. Gruending, T., Sauerland, V., Barahona, C., Herz, C., Nitsch, U.: Polyalanine - a practical polypeptide mass calibration standard for matrix-assisted laser desorption/ionization mass spectrometry and tandem mass spectrometry in positive and negative mode. *Rapid Commun Mass Spectrom.* 30, 681-683 (2016)
4. Fagerquist, C.K., Sultan, O.: A new calibrant for matrix-assisted laser desorption/ionization time-of-flight-time-of-flight post-source decay tandem mass spectrometry of non-digested proteins for top-down proteomic analysis. *Rapid Commun Mass Spectrom.* 26, 1241-1248 (2012)
5. Fagerquist, C.K., Garbus, B.R., Miller, W.G., Williams, K.E., Yee, E., Bates, A.H., Boyle, S., Harden, L.A., Cooley, M.B., Mandrell, R.E.: Rapid identification of protein biomarkers of *Escherichia coli* O157:H7 by matrix-assisted laser desorption ionization-time-of-flight-time-of-flight mass spectrometry and top-down proteomics. *Anal Chem.* 82, 2717-2725 (2010)
6. Fagerquist, C.K., Zaragoza, W.J.: *Proceedings of the 64th Conference on Mass Spectrometry and Allied Topics* (June 5-9, 2016, San Antonio, TX). Citation ID: 280941 MP 719.

7. Blattner, F.R., Plunkett, G. 3rd, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B., Shao, Y.: The complete genome sequence of *Escherichia coli* K-12. *Science*. 277, 1453-1462 (1997)
8. Cherepanov, P.P., Wackernagel, W.: Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene*. 158, 9-14 (1995)
9. Datsenko, K.A., Wanner, B.L.: One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6640-6645 (2000)
10. Guzman, L.M., Belin, D., Carson, M.J., Beckwith, J.: Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol.* 177, 4121-4130 (1995)
11. 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.* 75, 4341-4353 (2009)
12. Fagerquist, C.K., Sultan, O., Carter, M.Q.: Possible evidence of amide bond formation between sinapinic acid and lysine-containing bacterial proteins by matrix-assisted laser desorption/ionization (MALDI) at 355 nm. *J Am Soc Mass Spectrom.* 23, 2102-2114 (2012)

13. Chendo, C., Phan, T.N.T., Marque, S. R.A., Gimes, D., Laurence, C.: Mass spectrometry of nitroxide-terminated poly(4-vinylpyridine): A case of unwanted reactive MALDI. *Intl. J. Mass Spectrom.* 405, 50–58 (2016)
14. Spengler, B.: Post-source decay analysis in matrix-assisted laser desorption/ionization mass spectrometry of biomolecules. *J Mass Spectrom.* 32, 1019-1036 (1997)
15. Yu, W., Vath, J.E., Huberty, M.C., Martin, S.A.: Identification of the facile gas phase cleavage of the Asp-Pro and Asp-Xxx peptide bonds in matrix assisted laser desorption time-of-flight mass spectrometry. *Anal Chem.* 65, 3015–3023 (1993)
16. Paizs, B., Suhai, S.: Fragmentation pathways of protonated peptides. *Mass Spectrom Rev.* 24, 508-548 (2005)
17. Hu, Y., Hadas, B., Davidovitz, M., Balta, B., Lifshitz, C. Does IVR take place prior to peptide ion dissociation? *J Phys Chem A.* 107, 6507-6514 (2003)
18. Schlag EW, Selzle HL, Schanen P, Weinkauff R, Levine RD. Dissociation kinetics of peptide ions. *J Phys Chem A.* 110, 8497-8500 (2006)

Figure captions

Figure 1A shows the MS data of the bacterial cell lysate of pBAD18_*yahO*_EDL933_K-12_Δ*yahO* grown overnight on LBA supplemented with arabinose. **Figure 1B** shows an expanded m/z range of **Figure 1A**. **Figure 1C** is the MS/MS-PSD of the YahO protein ion at m/z 7711 (TIS: ± 100 Da). Fragment ions are identified by their m/z , ion type and number. The mature protein sequence (signal peptide removed) of YahO is shown between **Figures 1B** and **1C**. An asterisk indicates the site of cleavage and the resulting fragment ions are identified above (y-type) and below (b-type) the sequence. Methionine residues are underlined.

Figure 2 shows the ratio of the intensities of the complementary fragment ion pair [$y_{(70-x)}/b_x$] (including neutral losses) resulting from the “aspartic acid effect” (left Y-axis). Average and standard deviation of three technical replicates. Also plotted (right Y-axis) is the ratio of the number of protonation sites of each complementary fragment ion pair.

Figure 3 shows the total fragment ion intensity (including neutral losses) for each cleavage site of the polypeptide backbone of YahO. The fragment ion intensity of each cleavage site is normalized with respect to the total fragment ion intensity of the entire chain. Average and standard deviation of three technical replicates.

Figure 4A shows the MS data of the bacterial cell lysate of *E. coli* O157:H7 strain EDL933 grown overnight on LBA supplemented with ciprofloxacin. **Figure 4B** shows an expanded m/z range of **Figure 4A**. The peak at m/z 8027 is the matrix adducted B-subunit. **Figure 4C** is the MS/MS-PSD of the disulfide-reduced B-subunit at m/z 7819 (TIS: ± 100 Da). Fragment ions are identified by their m/z , ion type and number. The mature protein sequence (signal peptide removed) of B-subunit is shown between **Figures 4B** and **4C**. An asterisk indicates the site of

cleavage and the resulting fragment ions are identified above (y-type) and below (b-type) the sequence. Cysteine residues and the one methionine residues are underlined.

Figure 5A shows the MS data of the co-spotted bacterial cell lysates of *E. coli* O157:H7 strain EDL933 grown overnight on LBA supplemented with ciprofloxacin and pBAD18_*yahO*_EDL933_K-12_*ΔyahO* grown overnight on LBA supplemented with arabinose. The expanded m/z insert clearly shows the YahO and B-subunit ions. **Figure 5B** shows MS/MS-PSD of the YahO protein ion using a TIS window of -100 Da/ $+50$ Da. Fragment ions are identified by their m/z , ion type and number. The mature protein sequence (signal peptide removed) of YahO is shown between **Figures 5A** and **5B**. **Figure 5C** shows MS/MS-PSD of the B-subunit protein ion using a TIS window: -40 Da/ $+75$ Da.

Table 1. Strains used in this study

<i>E. coli</i> strains	Relevant Characteristics	Source or reference
RM1272	ATCC# 43895, EDL-933 <i>stx2a</i> isolate associated with hemorrhagic colitis outbreak. Originally isolated from raw hamburger.	Clay Center, Nebraska
MG1655 K-12	Genotype: F- <i>lambda</i> - <i>ilvG</i> - <i>rfb</i> -50 <i>rph</i> -1, serotype: OR:H48:K-	Blattner et al. 1997
MG1655 K-12/pCP20	Contains FLP recombinase, used for the excision of pKD4 derived kanamycin cassette from the K-12 genome	Cherepanov et al. 1995
MG1655 K-12/pKD4	Template plasmid for gene disruption with kanamycin resistance. The resistant gene is flanked by FRT sites	Datsenko and Warner, 2000
MG1655 K-12/pKD46	Temperature sensitive replication (repA101ts); encodes lambda Red genes (<i>exo</i> , <i>bet</i> , <i>gam</i>); native terminator (tL3) after <i>exo</i> gene; arabinose-inducible promoter for expression (ParaB); encodes <i>araC</i> for repression of ParaB promoter; Ampicillin resistant.	Datsenko and Warner, 2000
MG1655 K-12/pBAD18	Contains the arabinose inducible PBAD promoter. Used for construction of plasmid pWZ19, containing <i>yahO</i> from EDL933	Guzman et al., 1995
WJZ18	K-12 with kanamycin insert at the <i>yahO</i> loci, still bears pKD46	This study
WJZ19	K-12 with kanamycin insert at the <i>yahO</i> loci, cured of pKD46	This study
WJZ20	K-12 Δ <i>yahO</i> : <i>yahO</i> loci removed and kanamycin removed by pCP20. PCR verified	This study
pWZ19	pBAD18 vector containing the <i>bla</i> ampicillin/carbapenem resistant gene and subcloned <i>yahO</i> from EDL-933	This study

Table 2. Primers and probes used in this study

Name	Oligonucleotide Sequence 5'-3'	Genomic coordinates	Additional Information
WZF-EDL933- <i>yahO</i>	GATCGGTACCGTGGGTGATAGTGTGATA	Binds to 403138..403495 of accession NC_002655	Bold bases are the KpnI restriction site, used for cloning <i>yahO</i> from EDL933
WZR-EDL933- <i>yahO</i>	GATCAAGCTTGGGGATGTTACATAGGA	Binds to 403468..403495 of accession NC_002655	Bold bases are the HindIII restriction sites, used for cloning <i>yahO</i> from EDL933
WZF-P1- <i>yahO</i>	AATAATGTGGGTGGGTGATAGTGATAACAACCTCTGG AGCCGTAATATG GTGTAGGCTGGAGCTGCTTC	Binds to 346437..346486 of accession NC_000913	Bold bases are the priming sites for the pKD4 plasmid
WZR-P2- <i>yahO</i>	ATCACCGATAAACTTACGCATCGGAGATGTTTACATAG GATTCAGAATT AATGGGAATTAGCCATGGTCC	Binds to 346757..346806 of accession NC_000913	Bold bases are the priming sites for the pKD4 plasmid
WZF- <i>yahO</i> -DelConf	AGGTAAACGGCATGCAAA	Binds to 346273..346290 of accession NC_000913	Used to confirm deletion of <i>yahO</i> from MG1655
WZR- <i>yahO</i> -DelConf	AATATGATGGAACGACTGG	Binds to 347069..347087 of accession NC_000913	Used to confirm deletion of <i>yahO</i> from MG1655

Table 3: Calibrant fragment ions of YahO

Ion type/number	Average <i>m/z</i>
y13	1516.784
y23	2530.927
y28	3031.436
b32	3565.912
y34	3700.271
b36	4009.369
y38	4143.728
b42	4678.203
b47	5178.712
y49	5280.917
b57	6192.856
b65	7029.758

Table 4: Prominent fragment ions of B-subunit of Stx2

Ion type/number	Y₄₆	Y₅₃	Y₅₄
Theoretical average <i>m/z</i>	5112.666	5903.581	6018.669
Adjacent spots			
<i>m/z</i> (Ave ± SD)	5112.3 ± 0.3	5903.3 ± 0.0	6018.4 ± 0.3
Intensity (Ave ± SD)	14885 ± 5050	5642 ± 2399	7092 ± 2931
Same spot			
<i>m/z</i> (Ave ± SD)	5112.2 ± 0.1	5903.2 ± 0.0	6018.5 ± 0.2
Intensity (Ave ± SD)	5293 ± 861	2214 ± 473	2580 ± 569
Adjacent spots			
<i>m/z</i> (Ave ± SD)	5112.2 ± 0.2	5903.3 ± 0.2	6018.3 ± 0.2
Intensity (Ave ± SD)	40313 ± 4117	12118 ± 923	12874 ± 871
Same spot			
<i>m/z</i> (Ave ± SD)	5112.3 ± 0.1	5903.3 ± 0.2	6018.5 ± 0.2
Intensity (Ave ± SD)	13011 ± 2048	3712 ± 598	3629 ± 639

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5