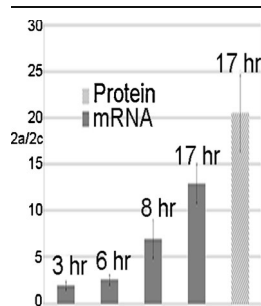


## RESEARCH ARTICLE

# Shiga Toxin 2 Subtypes of Enterohemorrhagic *E. coli* O157:H- E32511 Analyzed by RT-qPCR and Top-Down Proteomics Using MALDI-TOF-TOF-MS

Clifton K. Fagerquist, William J. Zaragoza

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



**Abstract.** We have measured the relative abundance of the B-subunits and mRNA transcripts of two Stx2 subtypes present in Shiga toxin-producing *Escherichia coli* (STEC) O157:H- strain E32511 using matrix-assisted laser desorption/ionization time-of-flight-time-of-flight tandem mass spectrometry (MALDI-TOF-TOF-MS/MS) with post source decay (PSD) and real time-quantitative polymerase chain reaction (RT-qPCR). Stx2a and Stx2c in STEC strain E32511 were quantified from the integrated peak area of their singly charged disulfide-intact B-subunit ions at  $m/z$  ~7819 and  $m/z$  ~7774, respectively. We found that the Stx2a subtype was 21-fold more abundant than the Stx2c subtype. The two amino acid substitutions (16D ↔ 16 N and 24D ↔ 24A) that distinguish Stx2a from Stx2c not only result in a mass

difference of 45 Da between their respective B-subunits but also result in distinctly different fragmentation channels by MS/MS-PSD because both substitutions involve an aspartic acid (D) residue. Importantly, these two substitutions have also been linked to differences in subtype toxicity. We measured the relative abundances of mRNA transcripts using RT-qPCR and determined that the *stx2a* transcript is 13-fold more abundant than *stx2c* transcript. In silico secondary structure analysis of the full mRNA operons of *stx2a* and *stx2c* suggest that transcript structural differences may also contribute to a relative increase of Stx2a over Stx2c. In consequence, toxin expression may be under both transcriptional and post-transcriptional control.

**Keywords:** Shiga toxin, *Escherichia coli*, MALDI-TOF-TOF, Top-down proteomics, RT-qPCR, mRNA, Bacteriophage, Antibiotic induction, Food safety

Received: 13 November 2014/Revised: 7 January 2015/Accepted: 7 January 2015/Published online: 10 February 2015

## Introduction

Shiga toxin-producing *Escherichia coli* (STEC) are increasingly linked to outbreaks of foodborne illness worldwide [1–3]. There are two Shiga toxin (Stx) types, Stx1 and Stx2. These types are further subdivided into subtypes: Stx1a, c, d, and Stx2a through g. A recent multicenter study proposed a systematic nomenclature for classifying Stx types and subtype [4].

Shiga toxin is an AB<sub>5</sub> toxin with a catalytically active A-subunit and five identical B-subunits. The five B-subunits form

a torus-shaped, noncovalent structure, which has 5-fold symmetry. Part of the A-subunit that includes the C-terminus (A2 fragment) fits within the “donut” hole of the B<sub>5</sub> complex, giving it greater stability. It is this quaternary complex that is responsible for transport of the catalytically-active part of the A-subunit (A1 fragment). The A2/B<sub>5</sub> complex is also responsible for recognizing and attaching to the glycolipid surface receptors on eukaryotic cells, specifically the glycolipid globotriaosylceramide (Gb3) receptor [5].

Amino acid (AA) sequence variations across Stx types and subtypes can result in differences in toxicity [6–9]. The AA variability is due primarily to substitutions that occur in the A2 fragment and the B-subunit. In contrast, there are fewer AA substitutions across subtypes in the catalytically-active A1 fragment of the A-subunit. This suggests that differences in the toxicity of Stx1, Stx2, and their subtypes have less to do with differences in A1 toxicity, whose sequence is highly

**Electronic supplementary material** The online version of this article (doi:10.1007/s13361-015-1076-3) contains supplementary material, which is available to authorized users.

Correspondence to: Clifton Fagerquist; e-mail: clifton.fagerquist@ars.usda.gov

conserved across subtypes, and more to do with stability of the AB<sub>5</sub> holotoxin, its transport, its attachment to surface receptors on eukaryotic cells, and its intracellular translocation.

It is not uncommon for STEC strains to have multiple *stx* genes resulting in multiple Stx types and/or subtypes. For example, *Escherichia coli* O157:H7 strain EDL933 has a *stx1* and *stx2* genes (both “a” subtypes) [10]. *E. coli* O91:H21 strain B2F1 has two nearly identical *stx2* genes (both “d” subtypes): one *stx2d* is located in a bacteriophage and the other *stx2d* is located in a cryptic phage [11, 12]. *E. coli* O157:H- strain E32511, which originates from a serious hemolytic uremic syndrome (HUS) infection, possesses two *stx2* gene subtypes: “a” and “c” [13–17].

A number of techniques are available for detection of *stx* genes or Stx toxin. Polymerase chain reaction (PCR) has made significant advances as a diagnostic tool for identification of the bacterial virulence genes including *stx* in pathogenic *E. coli*. PCR has been used to distinguish between *stx1* and *stx2* as well as distinguishing their subtypes [4]. This approach relies on developing internal primers of the target gene that are specific for the type and subtype. Success is dependent on the specificity of such primers to discriminate between highly similar subtype *stx* genes. Currently, there are over 90 STEC strains having distinct *stx2* genes and Stx2 amino acid sequences. Attempts have been made to logically group these sequences, delineate subtypes, and adopt a systematic nomenclature for subtyping [4]. However, as the variations in *stx2* genes and Stx2 sequences display an almost continuous variation from one subtype to the next, clear delineation between adjacent subtypes, especially for clinical subtypes, can be challenging. In addition, when a single strain has two genes or subtypes, primer amplification may indicate the presence of multiple *stx* gene or subtypes, or it may indicate that a single *stx* gene amplifies both primer subtypes because of shared homology between subtypes. The cost of PCR equipment is within the budgets of most research and regulatory labs, and PCR remains a common laboratory technique. A disadvantage of PCR is that the presence of a gene does not indicate whether (or by how much) a gene is actually expressed. As *stx* genes are bacteriophage-encoded, they are acquired by horizontal gene transfer from a bacteriophage, and Stx expression is largely dependent on phage replication within the bacteria [18].

Enzyme-linked immunosorbent assay (ELISA) has been shown to be quite useful for detection and quantitation of Stx and other bacterial toxins [19, 20]. A successfully developed and validated assay can be manufactured as a kit that can be made available commercially. As a kit, an ELISA is portable and readily used in the field, in a processing facility, or a testing laboratory. As such, it has been often used for conducting surveys for the presence of a target analyte in a commodity. A positive result by ELISA is considered *presumptive* positive until subsequent testing can confirm the result. Follow-up testing is usually performed using an orthogonal technique (e.g., mass spectrometry) in a laboratory. The disadvantage of ELISA is the possibility of cross-reactivity of the monoclonal antibody (mAb) (i.e., false positive) attributed to a lack of

specificity of the assay. In the case of a protein toxin, the mAb detects only a small surface region (i.e., epitope) corresponding to only a few amino acid residues that is unique to the protein toxin. The ability of the mAb to detect the target epitope, and *only* the target epitope, in a highly complex mixture reflects the dual analytical goals of sensitivity and selectivity. Sensitivity and selectivity are often mutually exclusive in ELISA/mAb development with an increase in selectivity often resulting in a loss of sensitivity and vice versa. For example, an ELISA has been shown to sensitively distinguish between Stx1 and Stx2 but it cannot distinguish between subtypes [19, 20]. Certain subtypes of Stx2 are more clinically relevant (a–d) than others (e–g). In consequence, definitive subtype identification is a very important analytical goal.

Another hybrid technique, immuno-PCR, combines elements of both mAbs and PCR to sensitively detect Stx [21]. Stx can also be detected using a bioassay that measures the toxicity in mammalian cells. Stx toxicity stems from the disruption of the eukaryotic ribosomal machinery of protein synthesis by the catalytically-active A1 fragment of the A-subunit [18]. A Vero cell-based assay can be used to monitor the inhibition of protein synthesis using a modified green fluorescent protein [22, 23]. A drawback of this approach is the time, labor, and facilities required to prepare and culture Vero cells (a mammalian cell line).

Mass spectrometry (MS)-based proteomics has several advantages for identification of protein toxins. First, it exploits the genetic and genomic data provided by gene sequencing and whole genome sequencing. Second, MS analysis is a chemical characterization based on molecular mass. A mass spectrometer with sufficient resolution and mass accuracy can confirm the elemental composition of a molecule. Third, tandem mass spectrometry (MS/MS) can provide structural and/or sequence-specific information and has found wide utility in proteomic sequencing as well as identification of post-translational modifications (PTM). Fourth, MS and MS/MS are not only sensitive but a highly specific techniques allowing *definitive* analyte identification in contrast with other techniques that lack specificity to distinguish Stx subtypes resulting in ambiguous or contradictory results. The initial cost of mass spectrometry instrumentation, e.g., matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF or MALDI-TOF-TOF) can be significant: \$200 to \$400 k. However, cost savings come from a simple sample preparation and high sample throughput. The rate-limiting step of the analysis is bacterial culturing. However, PCR, ELISA, and Vero cell-based assays often rely upon bacterial culturing as an important enrichment step.

Conventional proteomic identification involves enzymatic digestion (with trypsin) of the protein (or proteins) followed by MS and MS/MS of the resulting peptides. This approach is referred to as *bottom-up* proteomic analysis. In contrast, analysis of the intact, mature protein (or proteins) by MS and MS/MS *without* prior digestion is called *top-down* proteomic analysis. Both bottom-up and top-down have certain advantages and disadvantages and the approach taken is often dependent on the instrumentation and personnel available and the

objectives of the research. We have previously demonstrated a rapid technique for identification of protein biomarkers from unfractionated bacterial cell lysates using matrix-assisted laser desorption/ionization time-of-flight-time-of-flight mass spectrometry (MALDI-TOF-TOF-MS) and top-down proteomic analysis [24–26]. We also demonstrated this approach in the identification of the B-subunit and the A2 fragment of the A-subunit of Stx2 in *E. coli* O157:H7 strain EDL933 by antibiotic induction [27, 28]. In a subsequent study, we analyzed 26 STEC strains and were able to identify and distinguish between Stx2 subtypes: a, c, d, f, and g [29]. We observed significant variation in induced levels of Stx2 expression in these strains.

Reverse transcription-quantitative PCR (RT-qPCR) has been used by previous researchers to measure the mRNA transcript levels of *stx* [30, 31] and is increasingly used in testing laboratories as an indirect measurement of Stx production. As noted earlier, *E. coli* O157:H- strain E32511 has two Stx2 subtypes and is a highly studied clinical isolate known to cause serious HUS infection [13–17]. In order to better understand the factors responsible for production of Stx2 subtypes in *E. coli* O157:H- strain E32511, we measured the relative abundances of its two Stx2 subtypes from their mRNA transcripts and B-subunit proteins analyzed by real time-quantitative polymerase chain reaction (RT-qPCR) and MALDI-TOF-TOF-MS, respectively. In addition, we demonstrate the advantages of MALDI-TOF-TOF-MS/MS and post-source decay (PSD) for definitive top-down proteomic identification of multiple Stx2 subtypes expressed in a single bacterial strain. Portions of this work were presented at the 62nd Conference on Mass Spectrometry and Allied Topics (June 15–19, 2014, Baltimore, MD) [32].

## Materials and Methods

### Strains

**Cautionary Note** Shiga toxin-producing *E. coli* (STEC) are BSL-2 level microorganisms. All microbiological manipulations were conducted in a Class II biological safety cabinet. The strains used in this study are shown in Table 1. *E. coli* O157:H7 strain EDL933 (RM1272) is a STEC that has a *stx1* and a *stx2* gene. The *stx2* gene has been designated as a Stx2a subtype [4]. *E. coli* O157:H- strain E32511 (RM7004) is a STEC that has two *stx2* genes, one designated as a *stx2a* subtype and the other as a *stx2c* subtype [13–17]. *E. coli* strain RM7787 was isolated from a feral

pig in Northern California, and has a single *stx2* gene that is designated as a *stx2c* subtype [33].

### Culture Conditions

Bacterial cells were cultured from frozen stocks in Luria-Bertani broth (LBB) and incubated at 37°C with 250 rpm for 4 h. One hundred  $\mu$ L of cells were then plated on 1.5% Luria-Bertani agar (LBA) plates supplemented with 20 ng mL<sup>-1</sup> of ciprofloxacin, allowed to dry, and incubated statically at 37°C overnight. For time-course experiments, broth cultures were incubated and plated on solid media supplemented with 20 ng mL<sup>-1</sup> of ciprofloxacin as above. Plates were then incubated for 3, 6, 8, or 17 h at 37°C before being assayed.

### PCR for *stx2* Gene Sequencing

Initial amplification of *stx2* genes was carried out by performing gradient PCR on a Bio-Rad C1000 Touch Thermocycler (Bio-Rad, Hercules, CA, USA) from DNA derived directly from fresh *E. coli* colonies picked with a sterile pipette tip from the surface of solid agar plates added directly to PCR reaction mixtures. Primers were designed using CLC Main Workbench ver. 6.8.1 (CLCbio, Aarhus, Denmark) and purchased from Sigma-Aldrich (St. Louis, MO, USA). Each PCR reaction was carried out in a 50  $\mu$ L reaction mixture containing 41.7  $\mu$ L of molecular biology grade water (Fisher Scientific, Fair Lawn, NJ, USA), 5  $\mu$ L of 10x NEB Thermopol Reaction Buffer (New England Biolabs, Ipswich, MA, USA), 1  $\mu$ L of each primer (stock concentration 100  $\mu$ M), 1  $\mu$ L of dNTPs (10 mM), and 0.25  $\mu$ L of NEB Taq Polymerase (5000 U mL<sup>-1</sup>). Thermocycler settings were as follows: 10 min at 94°C followed by 29 cycles of 30 s at 94°C, 30 s of annealing, 2 min at 72°C, and a final extension time of 7 min at 72°C. The gradient of annealing temperatures were as follows: 58, 57.6, 56.8, 55.5, 53.6, 52.4, 51.5, and 51°C. Strains RM1272 (EDL933), RM7787, and RM7004 (E32511) were screened with all primer pairs to ensure specificity and to determine the optimal annealing temperature for each primer pair. PCR reactions were analyzed via electrophoresis on a 1% agarose gel (Fisher Scientific, Fair Lawn, NJ, USA). For large format gels, electrophoresis was carried out at 130 V for 70 min and for smaller analytical gels, electrophoresis was carried out at 85 V for 55 min. Agarose gels were analyzed using a Gel Doc XR+ imager with ImageLab software (Bio-Rad, Hercules, CA,

**Table 1.** Strains Used in This Study

Strains	Relevant characteristics	Source or reference
RM1272	ATCC# 43895, <i>E. coli</i> O157:H7 strain EDL-933. Stx2a subtype. Isolate associated with hemorrhagic colitis outbreak. Originally isolated from raw hamburger.	Clay Center, Nebraska.
RM7787	<i>E. coli</i> . Stx2c subtype.	Isolated from feral pig, California.
RM7004	<i>E. coli</i> O157:H- strain E32511. Stx2a and Stx2c subtypes. Enterohemorrhagic <i>E. coli</i> (EHEC) isolate associated with hemolytic uremic syndrome (HUS) in humans.	Tom Whittam, W02883.

USA). PCR products were analyzed with agarose gels using the illustra PCR DNA and Gel Band Purification Kit (GE Healthcare, Buckinghamshire, UK).

### *Stx2 Gene Sequencing*

Full *stx2* gene sequences were confirmed for strain RM1272 (EDL933), RM7004 (E32511), and RM7787 to ensure strain integrity and as a tool for probe and primer design. The same primers used in PCR were used for sequencing (Table 2). All PCR products were sequenced on an ABI3130xl (Applied Biosystems) sequencer using POP7 polymer (Applied Biosystems, Foster City, CA, USA) and BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems) following manufacturer specifications. The resulting sequences were assembled to either accession number X07865 or ECOSLTII-1 from NCBI using CLC Main Workbench v6.8.1 (CLCbio, Aarhus, Denmark).

### *RNA Extraction*

Total RNA extraction was carried out using the RNeasy Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's recommendations with the following modifications. RNAProtect Bacteria Reagent (Qiagen, Hilden, Germany) was mixed in a 2:1 ratio with fresh LB media. This mixture was dispensed onto plates and the cells were mixed with the stabilizer with a sterile plate spreader. This mixture was transferred to RNase-free microcentrifuge tubes, vortexed for 5 s, and incubated at ambient temperature for 5 min. This mixture was centrifuged for 5 min at 5000×g and the resulting pellet was used with the RNeasy Mini Kit. On-column DNase digestions were performed using the RNase-free DNase digestion kit (Qiagen) per manufacturer's instructions. RNA concentration and purity were determined with a Nanodrop 1000 (Thermo Scientific, Wilmington, DE, USA) and adjusted to a concentration of precisely 500 ng  $\mu\text{L}^{-1}$  to ensure qPCR reproducibility.

### *Reverse Transcription of mRNA to cDNA*

RNA concentrations were determined on a Nanodrop 1000. RNA samples were converted to cDNA using the iScript Reverse Transcription supermix (Bio-Rad, Hercules, CA, USA), which contains MMLV H<sup>+</sup> reverse transcriptase with RNase inhibitor, random primers, buffer, MgCl<sub>2</sub>, and stabilizers. The master mix for reverse transcription steps were as follows: 4  $\mu\text{L}$  of iScript super mix, 500 ng of RNA (variable volume), nuclease-free water (variable volume) to a total reaction volume of 20  $\mu\text{L}$ . The thermocycler protocol used was 5 min at 25°C, 30 min at 42°C, and 5 min at 85°C followed by a 4°C hold. Reactions were carried out on a C1000 Touch thermocycler (Bio-Rad). Purity and concentration were determined with a Nanodrop 1000.

**Table 2.** Primers and Probes Used in This Study

Name	Stx2 subtype	DNA sequence 5'-3'	Genomic coordinates	Additional information
WZX07865F1	a	TCCATTATCTGCATTATGC	55..73 of accession X07865	Used with WZX07865R2 to generate a 1610 bp fragment
WZX07865R2	a	GGTACCACCGAACCCTGTCTCGA	1643..1664 of accession X07865	Used with WZX07865F1 to generate a 1610 bp fragment or with WZX07865F to generate a 1664 bp fragment
WZX07865F	a	ATCCGATAGCTCATCGGAACAAGCT	1..25 of accession X07865	Used with WZX07865R2 to generate a 1664 bp fragment
WZECOSLTIIIF	c	GGTGGTGGTCCATTATCGGTATTC	3..26 of accession ECOSLTII-1	Used with WZECOSLTIIIR1 to generate a 1305 bp fragment
WZECOSLTIIIR1	c	ACTCTTTCCGCCACTTTTACTGTGAATG	1277..1306 of accession ECOSLTII-1	Used with WZECOSLTIIIF to generate a 1304 bp fragment
2a TaqMan probe	a	HEX-AGTGAAGGTTGACGGGAAAGA-BHQ1	1088..1108 of accession X07865	Binds to the coding region of the 2a B subunit
Sense primer	a	CGGATTGTGCTAAAAGGTA	1030..1047 of the coding region of X07865	Use with 2a antisense to form a 164 bp product
Antisense primer	a	GGTACTGGATTTCGATTGTGA	1174..1193 of the coding region of X07865	Use with 2a sense to form a 164 bp product
2c TaqMan probe	c	6FAM-AGTAAAAGTGCCCGAAAAGA-BHQ1	1088..1108 of the coding region of ECOSLTII	Binds to the coding region of the 2c B subunit
Sense primer	c	GGATTGCGCTAAAAGGTA	1227..1244 of the coding region of ECOSLTII	Use with 2c antisense to form a 163 bp product
Antisense primer	c	GGTACTGGATTTCGATTGTGA	1370..1389 of the coding region of ECOSLTII	Use with 2c sense to form a 163 bp product

### Probe Design for Quantitative PCR

Hydrolysis probes (TaqMan) were designed using CLC Main Workbench ver. 6.8.1 and purchased from Sigma-Aldrich (St. Louis, MO, USA). These probes exploit a 21-base pair region that contains five nucleotide base differences between the 2a and 2c subtypes of the B-subunit in RM7004 (E32511) (Table 2 and Figure 1). An analytical PCR was run with the flanking primers on the cDNA, performed to ensure that only one major product was generated.

### Probe Validation

Hydrolysis probes (TaqMan) were validated against RM1272 (EDL933) and RM7787, which contain one copy of the *stx2a* and *stx2c* genes, respectively. Reaction efficiency was determined using RNA extractions from RM7004 (E32511). The mRNA was extracted and converted into cDNA as described above. A six-point 10-fold serial dilution series was performed and assayed for each probe in triplicate over four biological replicates. The slope of the fitted line of the dilution series was used to calculate the reaction efficiency for both *stx2a* and *stx2c* probes using the MXPro software utilizing the formula:  $E = 10^{(1/s) - 1}$ , where  $E$  is the reaction efficiency and  $s$  is the slope of the fitted line. A two sample  $t$ -test was performed on the reaction efficiencies over four biological replicates.

### Probe Sensitivity

The sensitivity of the *stx2a* and *stx2c* probes were tested against oligonucleotides (Sigma-Aldrich) that precisely matched their respective probe targets in *E. coli* strain E32511. These target oligonucleotides were dissolved in TE buffer to a final concentration of 50  $\mu$ M. A six-point 10-fold dilution series was made with each target with the highest concentration at 200 pM and the lowest at 0.002 pM. The precise quantities of each oligonucleotide in their stock solutions were confirmed using the Nanodrop 1000 (Fisher Scientific). Nanodrop measurements and the sensitivity assay were repeated on 3 separate days with three technical replicates.

### Quantitative PCR

Quantitative PCR (qPCR) was performed using the Stratagene MXPro MX3000P thermocycler. SsoFast supermix was used

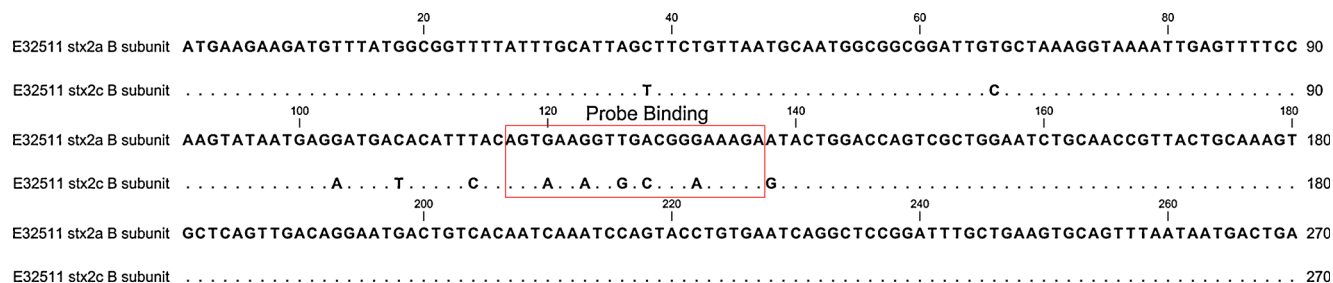
for mRNA qPCR (Bio-Rad, Hercules, CA, USA). The mastermix was prepared as follows (per reaction): 5  $\mu$ L of SsoFast supermix, 1  $\mu$ L of water, 250 nM of each primer, and 150 nM of the probe. For probe validation, a six-point 10-fold dilutions series was prepared for each extraction ranging from  $\approx 100$  ng  $\mu$ L<sup>-1</sup> to 0.001 ng  $\mu$ L<sup>-1</sup>. Each assay was performed with three technical replicates and three independent biological replicates. The qPCR protocol was 95°C for 10 min followed by 40 cycles of 95°C for 10 s, 55°C for 1 min, and 72°C for 1 min. A HEX dye (excite: 535 nm/emit: 555 nm) was used for the *stx2a* probe and a FAM dye (excite: 492 nm/emit: 516 nm) was used for the *stx2c* probe. The fluorescence threshold was set automatically by the Stratagene software (typically 10-fold higher than the background fluorescence). Due to the transient nature of mRNA transcript and the comparative stability of the Stx2 protein toxin, *stx2* mRNA levels were measured at 3, 6, 8, and 17 h. These experiments were conducted with at least four technical replicates and three biological replicates.

### RT-qPCR Data Analysis

Data analysis was carried out using the MXPro software (Stratagene, now Agilent). The significance was determined using a linear regression and coefficient determination ( $R^2$ ). Significance of the linear regression was analyzed by Tukey's  $t$ -test by testing the slope to be different from zero. Statistical analysis was performed using OriginLabs ver. 8.1 (Northampton, MA, USA).

### In Silico Secondary Structure Analysis of mRNA

Sequences were obtained from NCBI (accession numbers M76738, M59432, EF441603, and X07865) and confirmed via sequencing. Sequences were then uploaded to CLC Main Workbench 7.0.3 (<http://www.clcbio.com>) for analysis. For our analysis, we focused on the sequence from the start codon of the A-subunit to the stop codon of the B-subunit. Using CLC workbench, we computed 10 samples of suboptimal structures, calculated base pair probabilities, applied different energy rules for grossly asymmetric interior loops (GAIL), and included coaxial stacking energy rules using a minimum free energy approach (MFE). These structures were predicted using a modified version of the algorithm described by Rivas and Eddy



**Figure 1.** Nucleotide sequences of the B-subunit of *stx2a* and *stx2c* in *E. coli* O157:H- strain E32511. The hydrolysis probes bind to a 21-base pair region (boxed) that contains five nucleotide base difference between *stx2a* and *stx2c*

[34]. The thermodynamic energy parameters are used from the latest Mfold version 3 (<http://www.bioinfo.rpi.edu/~zukerm/ma/energy>).

### MALDI-TOF-TOF Mass Spectrometry and Top-Down Proteomic Analysis

Mass spectrometry (MS) and tandem mass spectrometry with post-source decay (MS/MS-PSD) data was collected using a 4800 MALDI-TOF-TOF mass spectrometer (AB SCIEX, Framingham, MA, USA) as described previously [24–26]. Briefly, cells were harvested from the agar plate after overnight culturing (17 h) using a sterile plastic 1.0  $\mu\text{L}$  microbiological loop, which when filled corresponded to approximately  $1 \times 10^9$  cells [24]. Cells were transferred to a screw-cap, O-ring lined microcentrifuge tube (Bartlesville, OK, USA) containing 300  $\mu\text{L}$  of extraction/lysis solution: 33% acetonitrile (Fisher Scientific), 67% water (Burdick & Jackson, Muskegon, MI, USA), and 0.2% trifluoroacetic acid (TFA, Sigma-Aldrich). All solvents were HPLC grade. The sample tube was tightly capped and bead-beat for 1 min followed by centrifugation at 16,000g for 2 min. One-half  $\mu\text{L}$  of supernatant was spotted onto multiple spots on a 384-spot stainless steel MALDI target and allowed to dry at room temperature. Each dried sample spot was then overlaid with 0.5  $\mu\text{L}$  of a 10 mg/mL solution of sinapinic acid (Protea Biosciences, WV, USA) dissolved in 33% acetonitrile, 67% water, and 0.2% TFA, and allowed to dry. Calibrants were also spotted onto the target for external instrument calibration in MS linear-mode (myoglobin, cytochrome-*c*, lysozyme) and MS/MS-PSD reflectron-mode (alkylated thioredoxin) [35]. The relative abundances of the disulfide-intact B-subunit of Stx2a and Stx2c subtypes were measured from the integrated area of their ion intensities in MS linear mode. Analysis was performed on four biological replicates (each with six technical replicates, i.e., spots) collected on four different days. Disulfide-reduced B-subunits were analyzed by MS/MS-PSD for definitive top-down proteomic identification [27–29]. Mass spectrometry analysis was not performed concurrently with RT-qPCR analysis.

## Results and Discussion

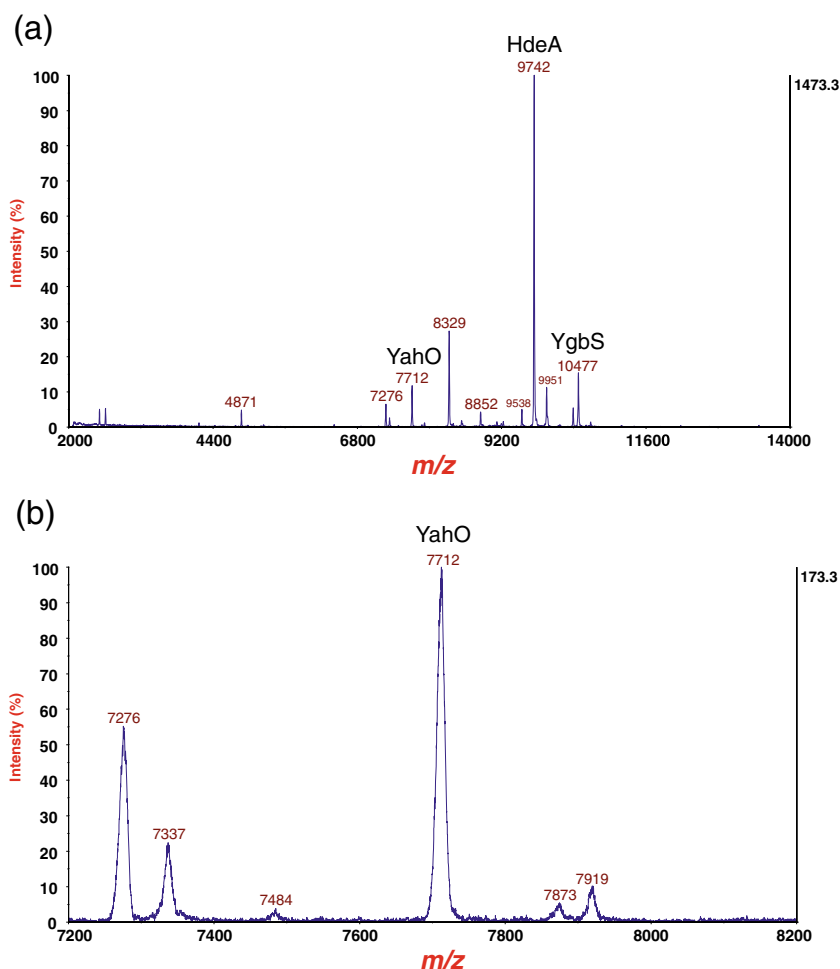
### Mass Spectrometry and Top-Down Proteomic Analysis

Figure 2a shows the MS spectrum of the bacterial cell lysate of *E. coli* O157:H- strain E32511 after overnight culturing (17 h) on LBA. A number of prominent peaks are observed, which correspond to high copy, cytosolic, or periplasmic proteins. For example, the peaks at  $m/z$  7712, 9742, and 10477 have been identified in other *E. coli* strains as the YahO protein, the acid stress protein HdeA, and the homeobox protein YbgS, respectively [26]. Figure 2b shows an expanded mass range from Figure 2a. The B-subunits of Stx2a or Stx2c were not detected in samples grown on LBA.

Figure 3a shows MS spectrum of the disulfide-reduced bacterial cell lysate of *E. coli* O157:H- strain E32511 after overnight culturing on solid agar supplemented with 20 ng  $\text{mL}^{-1}$  of ciprofloxacin. Figure 3b shows an expanded mass range from Figure 3a. The most abundant peak in the spectrum is at  $m/z$  7819 ( $[\text{M}+\text{H}]^+$ ) that corresponds to the singly charged, mature B-subunit of Stx2a (single peptide removed), which has a theoretical average  $m/z$  of 7818.6. The doubly charged B-subunit of Stx2a is observed at  $m/z$  3910 ( $[\text{M}+2\text{H}]^{+2}$ ). The peak at  $m/z$  7774 ( $[\text{M}+\text{H}]^+$ ) corresponds to the singly charged, mature B-subunit of Stx2c that has a theoretical average  $m/z$  of 7773.6. The observed mass difference between the putative Stx2a and Stx2c peaks is 45 Da, which corresponds to the theoretical mass difference of 45.0 Da.

The identity of these protein ions was confirmed by tandem mass spectrometry MS/MS-PSD. Figure 4a is the MS/MS-PSD spectrum of the peak at  $m/z$  7819 shown in Figure 3. In order to minimize spillover from adjacent peaks, a deliberately narrow asymmetric selection window ( $-15 \text{ Da}/+25 \text{ Da}$ ) for the precursor ion was used for ion isolation. Fragment ions are labeled by their  $m/z$ , ion type, and number. An expanded mass range insert highlights some of the less abundant fragment ions. The amino acid sequence of the B-subunit of Stx2a is shown above the spectrum. PSD of low charge state protein ions (e.g., +1 or +2) are known to fragment on the C-terminal side of aspartic acid (D) and glutamic acid (E) residues and on the N-terminal side of proline (P) residues [36, 37]. The favorability of fragmentation at D- and E-residues is the result of transfer of their acidic protons from their side chains to the polypeptide backbone, resulting in a weakening of the amide bond [36, 37]. As expected, the most abundant fragment ions in Figure 4a ( $y_{46}$ ,  $y_{53}$ ,  $y_{54}$ ) are the result of polypeptide backbone fragmentation on the C-terminal side of D residues (except for those proximal to the N- or C-termini). Lesser abundant fragments ions are the result of fragmentation on the C-terminal side of E residues ( $y_{43}$ ,  $y_{55}$ ,  $y_{61}$ ,  $b_{57}$ ,  $b_{64}$ ) and occasionally lysine residues ( $y_{65}$ ). The USDA top-down analysis software confirmed the top identification as the B-subunit of Stx2a ( $P$ -value= $3.4 \times 10^{-13}$  and USDA score=54.17 [24, 38]).

Figure 4b is the MS/MS-PSD spectrum of the peak at  $m/z$  7774 shown in Figure 3. In order to minimize ion spillover from the much more abundant peak at  $m/z$  7819, a narrow asymmetric ion selection window ( $-50 \text{ Da}/+15 \text{ Da}$ ) was used for ion isolation. Fragment ions are labeled by their  $m/z$ , ion type, and number. The amino acid sequence of the B-subunit of Stx2c is shown above the spectrum. The most abundant fragment ion is at  $m/z$  5860.3 ( $y_{53}$ ), which is the result of polypeptide backbone on the C-terminal side of the *only* D residue that is not proximal to the N- or C-termini. Note that the  $y_{53}$  fragment ion for Stx2a is at  $m/z$  5904.9. The difference in mass (or  $m/z$ ) of the  $y_{53}$  fragment ion of Stx2a and Stx2c is 44.6 Da (or Th), which is close to the expected 44.0 Da mass difference that includes an amino acid substitution at residue 24: A  $\leftrightarrow$  D. Owing to the very low signal-to-noise (S/N) of this MS/MS-PSD spectrum, it was not possible to perform top-down analysis using our software. However, manual inspection of the

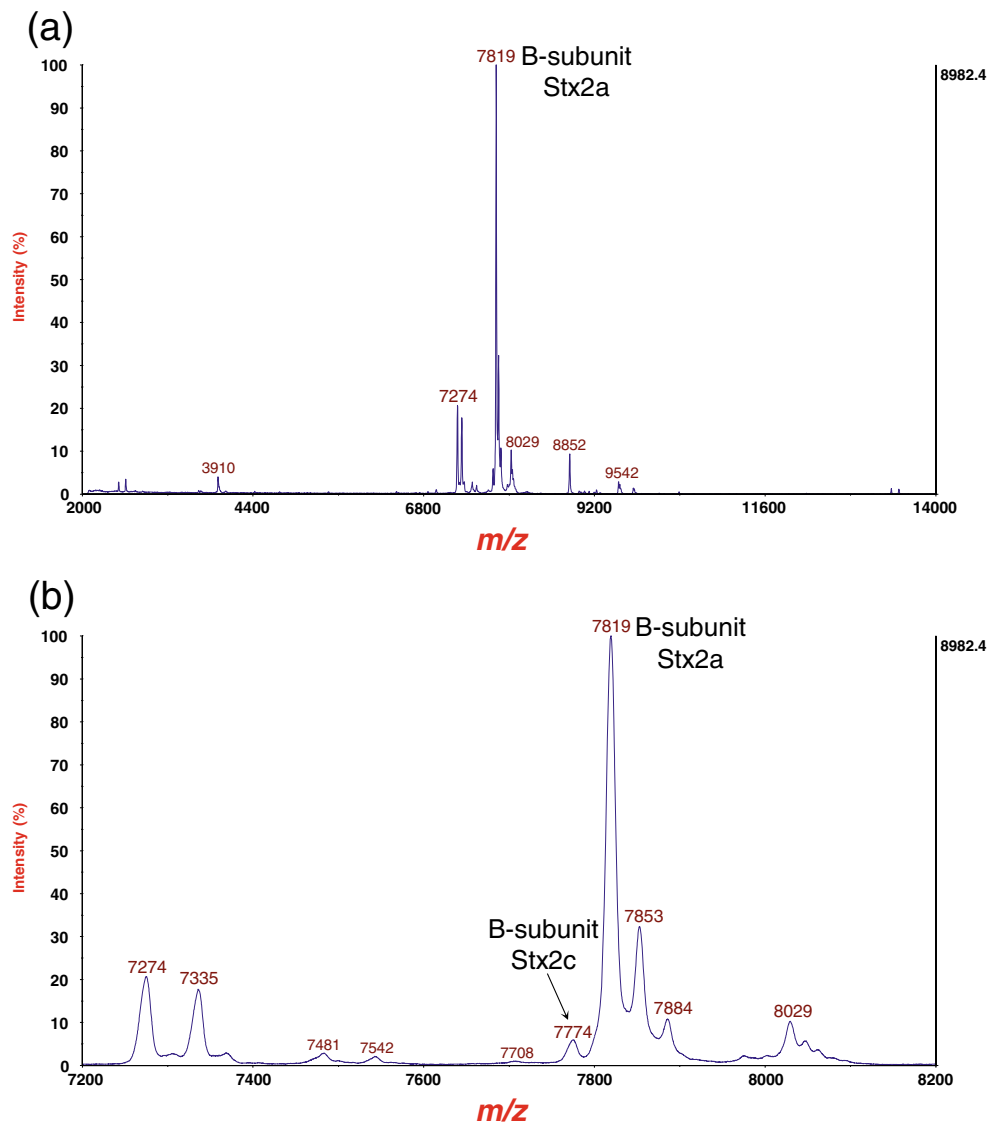


**Figure 2.** (a) MS spectrum of the bacterial cell lysate of *E. coli* O157:H- strain E32511 after overnight culturing on LBA. (b) Expanded mass range of (a)

MS/MS-PSD spectrum revealed (as expected) a prominent fragment ion ( $y_{53}$ ) resulting from fragmentation on the C-terminal side of the only D-residue not located at the ends of the polypeptide chain. Two other fragment ions were also identified. That, coupled with the precursor ion  $m/z$ , leaves no doubt as to the identity of this protein ion.

The poor quality of this MS/MS-PSD spectrum was due to two factors. First, the Stx2c precursor ion is ~20-fold less abundant than the Stx2a precursor ion. Obtaining a high quality MS/MS-PSD spectrum requires a strong precursor ion signal. For example, Supplemental Figure 1 shows the MS/MS-PSD of the Stx2c B-subunit of STEC strain RM7787, which was strongly induced on LBA supplemented with ciprofloxacin. As noted earlier, RM7787 has a single *stx2c* gene resulting in an amino acid sequence that is identical to Stx2c of STEC strain E32511. The stronger precursor ion signal results in a high quality MS/MS-PSD spectrum with a much improved S/N. Second, a narrow *asymmetric* selection window was used to eliminate any ion spillover from the more abundant of Stx2a precursor ion during MS/MS-PSD analysis of the Stx2c precursor ion. However, the narrowness of such a window invariably results in some loss of Stx2c ions as well.

As the B-subunits of Stx2a and Stx2c have nearly identical amino acid sequences, their ionization efficiencies by MALDI in positive ion mode would be expected to be the same given the fact that ionization (protonation) would most likely occur at a basic residue or the N-terminus, and there are no differences in the number (seven) or location of basic residues in these two proteins. Thus, the integrated peak areas of their respective ion intensities would reflect their relative abundance in the sample. The B-subunit ion intensities of Stx2a and Stx2c were measured after overnight culturing (17 h) in four biological replicates on different days and are shown in Figure 5. Stx2a is ~21-fold more abundant than Stx2c. The absence of detection of Stx2a and Stx2c when E32511 was grown on agar alone (Figure 2), and the detection of both Stx2a and Stx2c when E32511 was grown on medium supplemented with ciprofloxacin (Figure 3) strongly suggests that *stx2a* and *stx2c* genes are located in one (or more) *inducible* bacteriophages inserted into the bacterial genome. It is difficult to prove definitively how many bacteriophages are present in E32511 in the absence of a fully annotated genomic sequence. However, the disparate abundances of Stx2a and Stx2c may be due to their respective genes being located on two separate bacteriophage genomes



**Figure 3.** (a) MS spectrum of the bacterial cell lysate (disulfide-reduced) of *E. coli* O157:H- strain E32511 after overnight culturing on LBA supplemented with 20 ng mL<sup>-1</sup> of ciprofloxacin. (b) Expanded mass range from (a)

that are differentially induced by the antibiotic. Alternatively, a single inducible prophage containing both *stx2a* and *stx2c* genes is not impossible. The large differential production of Stx2a and Stx2c led us to measure the mRNA transcripts of *stx2a* and *stx2c* by RT-qPCR to further elucidate factors contributing to toxin production.

#### RT-qPCR: Probe Validation

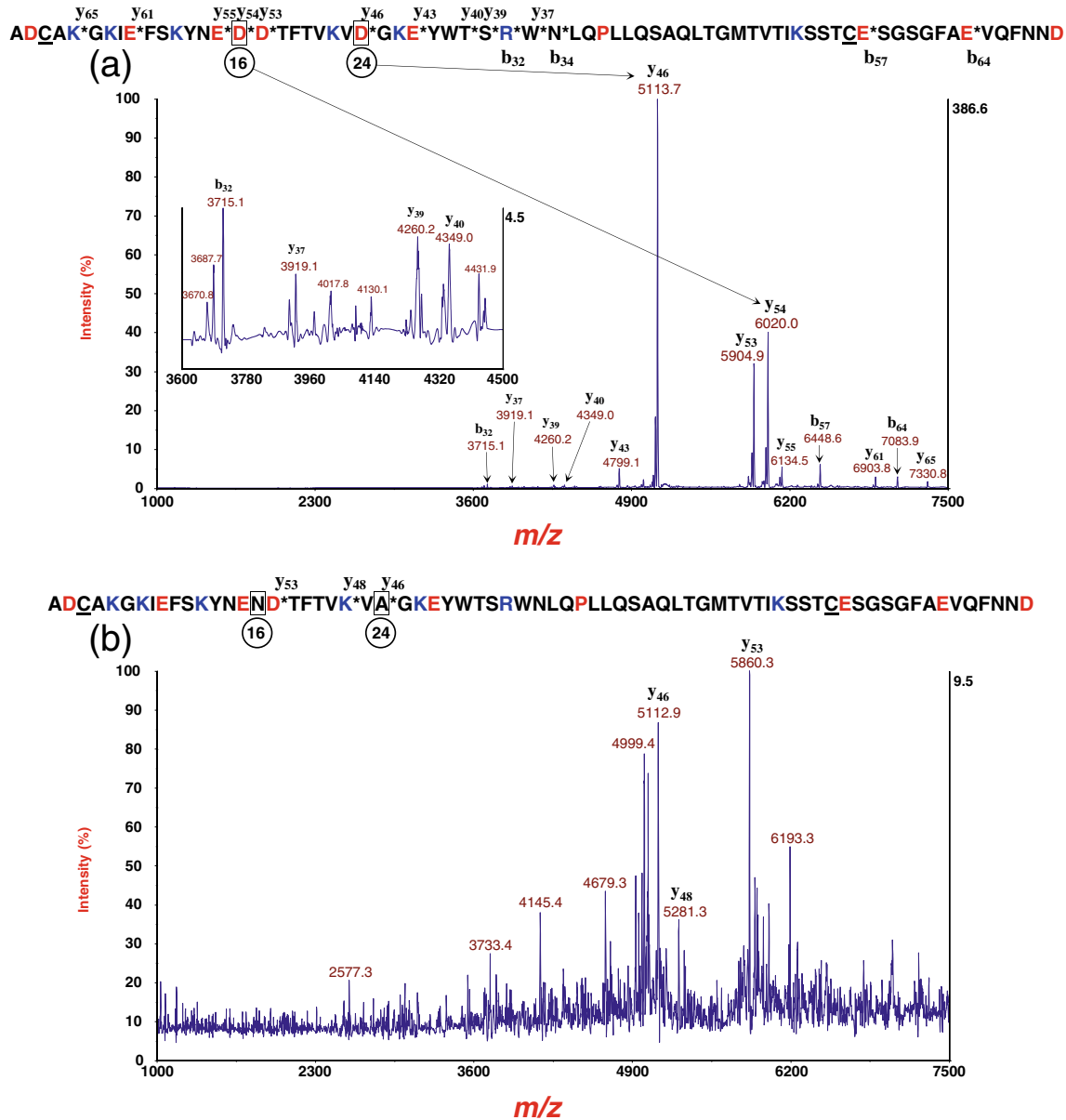
We used RM1272 and RM7787 as control strains (Table 1) to determine probe specificity. RM1272 (EDL933) has one copy of the *stx2a* subtype. RM7787 has one copy of the *stx2c* subtype. We determined that the *stx2a* probe gave no signal when tested against the total cDNA of RM7787, whereas a positive signal ( $C_t=25.75$ ) was obtained when tested against the total cDNA of RM1272. Similarly, the *stx2c* probe yielded a positive signal ( $C_t=24.95$ ) when tested against RM7787 and

no signal when assayed against RM1272. These results were confirmed similarly with the synthetic oligonucleotide targets.

#### RT-qPCR: Probe Reaction Efficiency

A six-point 10-fold serial dilution series was performed with each probe in triplicate over four biological replicates of RM7004 (E32511). The measured reaction efficiency incorporates the efficiency of the flanking primers to bind to their targets as well as the hydrolysis probes to bind to their respective targets. The average reaction efficiency for both the *stx2a* and *stx2c* probes over four biological replicates was  $95.8\% \pm 1.8\%$  and  $94.8\% \pm 1.3\%$ , respectively. A two sample *t*-test performed on the reaction efficiencies of the *stx2a* and *stx2c* probes at a 0.05 confidence level resulted in a *P*-value of 0.43229 that indicated no statistical difference in the reaction efficiencies of the *stx2a* and *stx2c* probes. In consequence, a





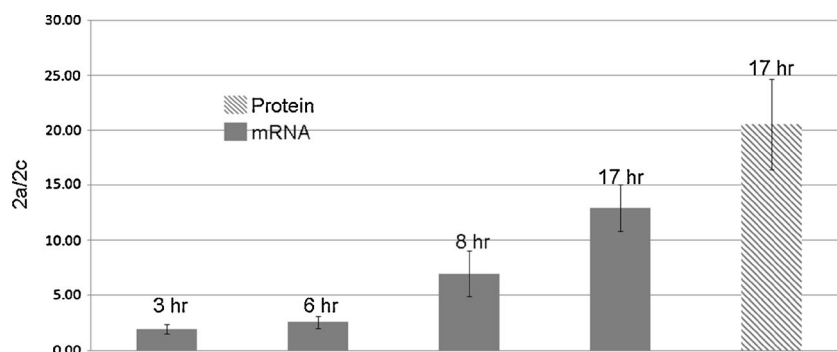
**Figure 4.** (a) MS/MS-PSD spectrum of the peak at  $m/z$  7819 shown in Fig. 3. The amino acid sequence of the B-subunit of Stx2a is shown above the spectrum. An asterisk (\*) indicates the site of fragmentation with the corresponding fragment ion type/number above (or below) the site of fragmentation. Boxed residues indicate an amino acid substitution that is different from that found in the Stx2c sequence. (b) MS/MS-PSD spectrum of the peak at  $m/z$  7774 shown in Fig. 3. An asterisk (\*) indicates the site of fragmentation with the corresponding fragment ion type/number above the site of fragmentation. Boxed residues indicate that a residue is different from that in the Stx2a sequence (a)

reaction efficiency of 95% was used on all subsequent calculations of mRNA abundances.

#### RT-qPCR: Probe Sensitivity

Hydrolysis probes were tested against synthetic oligonucleotides that precisely matched their targets in strain RM7004 (E32511). If the *stx2a* and *stx2c* probes have the same sensitivity to their respective targets, their  $C_t$  values should be the same by RT-qPCR if the initial amounts of each

oligonucleotide are the same. Fifty  $\mu\text{M}$  stock solutions of each oligonucleotide were prepared. Using the reported amounts provided by the vendor, the stock solutions were 2609  $\text{ng } \mu\text{L}^{-1}$  for the *stx2a* oligonucleotide and 2605  $\text{ng } \mu\text{L}^{-1}$  for the *stx2c* oligonucleotide. These stock solutions were then measured on the Nanodrop. The *stx2a* oligonucleotide concentration was found to be 2604  $\text{ng } \mu\text{L}^{-1}$ , whereas the *stx2c* oligonucleotide concentration was 2743  $\text{ng } \mu\text{L}^{-1}$ . The reported amounts provided by the vendor were guaranteed to have an error of not more than 10% (which they were). The Nanodrop



**Figure 5.** The relative abundance of mRNA transcript levels of *stx2a* vs. *stx2c* as measured by RT-qPCR at four time points: 3, 6, 8, 17 h of culturing *E. coli* O157:H- strain E32511 on solid agar supplemented with 20 ng mL<sup>-1</sup> of ciprofloxacin. The relative abundance of the B-subunits (disulfide-intact) of Stx2a and Stx2c measured after overnight culturing (17 h)

has a margin of error of 2%. On the basis of the Nanodrop measurements, there was 5.3% more *stx2c* oligonucleotide than *stx2a* oligonucleotide.

RT-qPCR measurements of a six-point 10-fold dilution series of the stock solutions revealed a  $C_t^{stx2a}/C_t^{stx2c}$  ratio of  $0.97 \pm 0.01$ . The  $C_t$  value is inversely proportional to the amount of initial target. For example, a lower  $C_t$  value indicates either more initial target or greater sensitivity of the probe to its target (or both). As there was 5.3% more *stx2c* oligonucleotide than *stx2a*, a value of 1.053 for  $C_t^{stx2a}/C_t^{stx2c}$  should have been obtained if their probe sensitivities were the same. However, the value of 0.97 indicates that the sensitivity of the *stx2a* probe to its target is ~8% greater than the sensitivity of the *stx2c* probe to its target.

The reaction efficiency was also monitored and found to closely match the calculated reaction efficiencies of the assay on RM7004 (E32511). The *stx2a* probe was found to have a reaction efficiency of 95.1% ( $R^2=0.998$ ) when tested against the oligonucleotide synthetic target. The *stx2c* probe had a reaction efficiency of 94.2% ( $R^2=0.990$ ).

### RT-qPCR

Figure 5 shows four time points (3, 6, 8, 17 h) for culturing *E. coli* O157:H- strain E32511 on solid agar supplemented with 20 ng mL<sup>-1</sup> of ciprofloxacin and measuring the ratio of mRNA transcript levels of *stx2a* versus *stx2c* by RT-qPCR. No correction was made for differences in probe sensitivity. An increase in the *stx2a/stx2c* ratio is observed with increasing culturing time. The *stx2a/stx2c* ratio at 17 h is less than the protein production ratio. Assuming complete correlation between mRNA transcript levels and protein production, one would expect that the 17 h mRNA transcript ratio would be higher than the protein production ratio (i.e., ~21-fold) by an amount sufficient to compensate for the lower transcript ratios at the 3, 6 and 8 h time points. Our results would suggest that the correlation between mRNA transcript levels and protein abundances appears to be present but is not strong. We conducted an in silico analysis of the mRNA transcripts of *stx2a* and *stx2c* to determine if secondary structure might contribute to differential protein production.

### Secondary Structural Analysis of mRNA of *stx2a* and *stx2c*

In silico secondary structure analysis of the mRNA transcripts of *stx2a* and *stx2c* indicate that the Gibbs free energy of the *stx2a* transcript is  $\Delta G = -399.2$  kcal mol<sup>-1</sup>, whereas the *stx2c* transcript is  $\Delta G = -389.4$  kcal mol<sup>-1</sup>. The thermodynamic stability of the *stx2a* transcript is greater than the *stx2c* transcript by about approximately  $-10$  kcal mol<sup>-1</sup>. In addition, the 5' untranslated regions (UTR) of the two B-subunit genes are distinct and give different predicted secondary structures.

Translation is regulated by events that control the formation of the elongation complex. Messenger RNA (mRNA) is the primary variable component in translation initiation among prokaryotes. The sequence and secondary structure of mRNA have effects on interactions with the translational machinery, thus influencing its frequency and efficiency. The pyrimidine-rich 5' UTR of canonical mRNAs plays a role in binding the S1 protein of the 30 s ribosomal initiation complex [39]. The S1 protein is known to alter mRNA secondary structure and is essential for bacteriophage replicase activity [40]. Translation initiation requires interaction between ribosomal protein S1 and the ribosomal binding site (RBS). These interactions occur at the local, single-stranded mRNA level [41]. Secondary structures can lower translational frequencies and efficiencies. Post-transcriptional control of mRNA appears to play a major role in controlling gene expression [42, 43]. Prokaryotic mRNA transcripts exhibit half-lives as short as a few min in vivo. mRNA templates determine transcript stability, which in turn determines the duration a transcript is available for translation before being targeted by the cellular RNA-degradation machinery. This complicates analysis of gene expression at the transcript level as an indirect measurement of protein translation.

### Top-Down Versus Bottom-Up Proteomic Analysis for Distinguishing Stx2 Subtypes

Top-down proteomic analysis has significant advantages over bottom-up analysis for identifying Stx2 subtypes [29, 44]. By top-down, the sequence-specific differences between similar or nearly identical protein sequences are reflected in (1) variations

in the  $m/z$  of the intact protein ions; (2) variations in  $m/z$  of the sequence-specific fragment ions; (3) variations in sequence-specific fragmentation channels. Unlike top-down, bottom-up analysis does *not* measure the  $m/z$  of the mature, intact protein ion, and the sequence coverage by bottom-up can vary considerably with 50% coverage being considered excellent. The gaps in bottom-up sequence coverage are caused by the non-detection and/or non-fragmentation of certain tryptic peptides generated from the protein digestion. In addition, it is more challenging for bottom-up to confirm PTMs such as N-terminal methionine or signal peptide removal because these PTMs can only be *inferred* from the absence of sequence coverage at that location. In contrast, top-down analysis can confirm such PTMs from both the  $m/z$  of the mature, intact protein ion and the  $m/z$  of sequence-specific fragment ions. For example, the two residue substitutions (16 and 24) that distinguish the mature B-subunits of Stx2a and Stx2c (as well as removal of their signal peptides) were confirmed from the  $m/z$  of their respective intact protein ions, the difference in  $m/z$  between the two protein ions (i.e., 45 Da) and differences in  $m/z$  of sequence-specific fragment ions. In addition, the two amino acid substitutions that distinguish the B-subunit sequence of Stx2a from Stx2c both involve an aspartic acid (D) residue, which not only result in differences in fragment ion  $m/z$  but also significantly alters the fragmentation channels of these protein ions. Fragmentation of the polypeptide backbone is highly favored on the C-terminal side of D-residues [36, 37] and this facile fragmentation channel is clearly demonstrated in Figure 4 and Supplementary Figure 1.

Most interestingly, the substitutions at residues 16 and 24 have also been linked to differences in subtype toxicity. For example, Lindgren et al. suggested that the aspartic acid at residue 16 plays a critical role in the binding of the Stx2a holotoxin to the glycolipid Gb3 receptor compared with the binding of the Stx2c holotoxin to the Gb3 receptor where residue 16 is an alanine [8]. They also noted that other researchers had implicated the aspartic acid at residue 24 as responsible for differences in holotoxin/receptor binding [9]. The issue of Stx2 holotoxin/Gb3 receptor binding is beyond the scope of the current work; however, native state mass spectrometry [45, 46], which involves analysis of higher order protein complexes by electrospray ionization mass spectrometry (ESI-MS), may be used to confirm the importance of residues 16 and 24 on holotoxin stability and Gb3 binding. In any case, MALDI-TOF-TOF-MS/MS-PSD is particularly well-suited for detecting the presence (or absence) of amino acid substitutions involving D-residues (and other residues) in a polypeptide sequence. Top-down proteomic analysis is better (than bottom-up) at detecting subtle differences in proteoform sequences in general and Stx types and subtypes in particular.

## Conclusions

We found that the B-subunit of Stx2a subtype was 21-fold more abundant than the B-subunit of the Stx2c subtype by

MALDI-TOF-TOF mass spectrometry. The two amino acid substitutions that distinguish Stx2a from Stx2c not only result in a measurable mass difference of 45 Da between their mature B-subunits but also result in unique fragmentation channels by MS/MS-PSD providing definitive identification and differentiation of these two subtypes. Importantly, these two amino acid substitutions have also been linked to differences in subtype toxicity. From an analytical perspective, MALDI-TOF-TOF-MS/MS-PSD is uniquely positioned to rapidly identify and differentiate these two critical substitutions that have been linked to Stx2 subtype toxicity. We also found that the *stx2a* mRNA transcript was 13-fold (12-fold if one corrects for differences in probe sensitivity) more abundant than the *stx2c* transcript by RT-qPCR. In silico secondary structure analysis suggests a greater thermodynamic stability of the *stx2a* transcript compared with *stx2c*, and there are structural differences in the 5' UTR between *stx2a* and *stx2c*. These mRNA transcript differences may contribute to a further increase of Stx2a over Stx2c suggesting that toxin production may be under both transcriptional and post-transcriptional control.

## Acknowledgments

This research was supported by USDA-ARS CRIS project 5325-42000-047-00D.

## References

1. Frank, C.D., Werber, J.P., Cramer, M., Askar, M., Faber, M., an der Heiden, H., Bernard, A., Fruth, R., Prager, A., Spode, M., Wadl, A., Zoufaly, S., Jordan, M.J., Kemper, P., Follin, L., Müller, L.A., King, B., Rosner, U., Buchholz, K., Stark, K., Krause, G.: Epidemic profile of Shiga-toxin-producing *Escherichia coli* O104:H4 outbreak in Germany. *N. Engl. J. Med.* **365**, 1771–1780 (2011)
2. Bielaszewska, M., Mellmann, A., Zhang, W., Köck, R., Fruth, A., Bauwens, A., Peters, G., Karch, H.: Characterisation of the *Escherichia coli* strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: a microbiological study. *Lancet Infect. Dis.* **11**, 671–678 (2011)
3. Brzuszkiewicz, E., Thürmer, A., Schuldes, J., Leimbach, A., Liesegang, H., Meyer, F.-D., Boelter, J., Petersen, H., Gottschalk, G., Daniel, R.: Genome sequence analyses of two isolates from the recent *Escherichia coli* outbreak in Germany reveal the emergence of a new pathotype: entero-aggregative-haemorrhagic *Escherichia coli* (EAHEC). *Arch. Microbiol.* **193**, 883–891 (2011)
4. Scheutz, F., Teel, L.D., Beutin, L., Piérard, D., Buvens, G., Karch, H., Mellmann, A., Caprioli, A., Tozzoli, R., Morabito, S., Strockbine, N.A., Melton-Celsa, A.R., Sanchez, M., Persson, S., O'Brien, A.D.: Multicenter evaluation of a sequence-based protocol for subtyping Shiga toxins and standardizing Stx nomenclature. *J. Clin. Microbiol.* **50**, 2951–2963 (2012)
5. Bunge, J.C., Melton-Celsa, A.R., O'Brien, A.D.: Shiga toxin type 2dact displays increased binding to globotriaosylceramide in vitro and increased lethality in mice after activation by elastase. *Toxins (Basel)* **5**, 2074–2092 (2013)
6. Fuller, C.A., Pellino, C.A., Flagler, M.J., Strasser, J.E., Weiss, A.A.: Shiga toxin subtypes display dramatic differences in potency. *Infect. Immun.* **79**, 1329–1337 (2011)
7. Yamasaki, S., Furutani, M., Ito, K., Igarashi, K., Nishibuchi, M., Takeda, Y.: Importance of arginine at position 170 of the A subunit of Vero toxin 1 produced by enterohemorrhagic *Escherichia coli* for toxin activity. *Microb. Pathog.* **11**, 1–9 (1991)
8. Lindgren, S.W., Samuel, J.E., Schmitt, C.K., O'Brien, A.D.: The specific activities of Shiga-like toxin type II (SLT-II) and SLT-II-related toxins of enterohemorrhagic *Escherichia coli* differ when measured by Vero cell cytotoxicity but not by mouse lethality. *Infect. Immun.* **62**, 623–631 (1994)

9. Ito, H., Nishibuchi, M., Takeda, Y.: Analysis of the antigenic difference between Vero toxin 2 (VT2) and VT2 variant (VT2vh) of Verotoxin-producing *Escherichia coli* by a site-directed mutagenesis. *FEMS Microbiol. Lett.* **63**, 27–30 (1991)
10. Perna, N.T., Plunkett, G., 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., Potamouisis, K.D., Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C., Welch, R.A., Blattner, F.R.: Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* **409**, 529–533 (2001)
11. Lindgren, S.W., Melton, A.R., O'Brien, A.D.: Virulence of enterohemorrhagic *Escherichia coli* O91:H21 clinical isolates in an orally infected mouse model. *Infect. Immun.* **61**, 3832–3842 (1993)
12. Ito, H., Terai, A., Kurazono, H.: Cloning and nucleotide sequencing of Vero toxin 2 variant genes from *Escherichia coli* O91: H21 isolated from a patient with the hemolytic uremic syndrome. *Microb. Pathog.* **8**, 47–60 (1990)
13. Scotland, S.M., Smith, H.R., Willshaw, G.A., Rowe, B.: Vero cytotoxin production in strain of *Escherichia coli* is determined by genes carried on bacteriophage. *Lancet* **2**, 216 (1983)
14. Willshaw, G.A., Smith, H.R., Scotland, S.M., Rowe, B.: Cloning of genes determining the production of vero cytotoxin by *Escherichia coli*. *J. Gen. Microbiol.* **131**, 3047–3053 (1985)
15. Dickie, N., Speirs, J.: Purification of an *Escherichia coli* serogroup O157: H7 verotoxin and its detection in North American hemorrhagic colitis isolates. *J. Clin. Microbiol.* **27**, 1973–1978 (1989)
16. Hii, J.H., Gyles, C., Moroooka, T., Karmali, M.A., Clarke, R., De Grandis, S., Brunton, J.L.: Development of verotoxin 2- and verotoxin 2 variant (VT2v)-specific oligonucleotide probes on the basis of the nucleotide sequence of the B cistron of VT2v from *Escherichia coli* E32511 and B2F1. *J. Clin. Microbiol.* **29**, 2704–2709 (1991)
17. Schmitt, C.K., Mckee, M.L.: O'Brien, A.D.: Two copies of Shiga-like toxin II-related genes common in enterohemorrhagic *Escherichia coli* strains are responsible for the antigenic heterogeneity of the O157: H- strain E32511. *Infect Immun.* **59**, 1065–1073 (1991)
18. Johannes, L., Römer, W.: Shiga toxins—from cell biology to biomedical applications. *Nat. Rev. Microbiol.* **8**, 105–116 (2010)
19. Skinner, C., Patfield, S., Stanker, L.H., Fratamico, P., He, X.: New high-affinity monoclonal antibodies against Shiga toxin 1 facilitate the detection of hybrid Stx1/Stx2 in vivo. *PLoS One* **9**, e99854 (2014)
20. He, X., Patfield, S., Hnasko, R., Rasooly, R., Mandrell, R.E.: A polyclonal antibody based immunoassay detects seven subtypes of Shiga toxin 2 produced by *Escherichia coli* in human and environmental samples. *PLoS One* **8**, e76368 (2013)
21. He, X., Qi, W., Quiñones, B., McMahon, S., Cooley, M., Mandrell, R.E.: Sensitive detection of Shiga toxin 2 and some of its variants in environmental samples by a novel immuno-PCR assay. *Appl. Environ. Microbiol.* **77**, 3558–3564 (2011)
22. Quiñones, B., Swimley, M.S.: Use of a Vero Cell-Based Fluorescent Assay to Assess Relative Toxicities of Shiga toxin 2 Subtypes from *Escherichia coli*. In: Holst, O., Holst, O. (eds.) *Methods in Molecular Biology, Microbial Toxins*, Vol. 739, pp. 61–71. Humana Press, Totowa (2011)
23. Massey, S., Quiñones, B., Teter, K.: A Cell-Based Fluorescent Assay to Detect the Activity of Shiga Toxin and Other Toxins that Inhibit Protein Synthesis. In: Holst, O. (ed.) *Methods in Molecular Biology, Microbial Toxins*, Vol. 739, pp. 49–59. Totowa, Humana Press (2011)
24. 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)
25. Fagerquist, C.K., Harden, L.A., Garbus, B.R. United States Patent. 8,160,819 B2. United States.
26. Fagerquist, C.K., Garbus, B.R., Miller, W.G., Williams, K.E., Yee, E., Bates, A.H., Harden, L.A., Cooley, M.B., Mandrell, R.E.: Rapid identification of protein biomarkers of *Escherichia coli* O157:H7 by matrix-assisted laser mass spectrometry and top-down proteomics. *Anal. Chem.* **82**, 2717–2725 (2010)
27. Fagerquist, C.K., Sultan, O.: Induction and identification of disulfide-intact and disulfide-reduced  $\beta$ -subunit of Shiga toxin 2 from *Escherichia coli* O157:H7 using MALDI-TOF-TOF-MS/MS and top-down proteomics. *Analyst* **136**, 1739–1746 (2011)
28. Fagerquist, C.K., Sultan, O.: Top-down proteomic identification of furin-cleaved  $\alpha$ -subunit of Shiga toxin 2 from *Escherichia coli* O157:H7 using MALDI-TOF-TOF-MS/MS. *J. Biomed. Biotechnol.* **2010**, Article ID:123460, 11 pages (2010)
29. Fagerquist, C.K., Zaragoza, W.J., Sultan, O., Woo, N., Quiñones, B., Cooley, M.B., Mandrell, R.E.: Top-down proteomic identification of Shiga toxin 2 subtypes from Shiga toxin-producing *Escherichia coli* by matrix-assisted laser desorption/ionization-tandem time of flight mass spectrometry. *Appl. Environ. Microbiol.* **80**, 2928–2940 (2014)
30. Zhang, W., Bielaszewska, M., Friedrich, A.W., Kuczius, T., Karch, H.: Transcriptional Analysis of genes encoding Shiga toxin 2 and its variants in *Escherichia coli*. *Appl. Environ. Microbiol.* **71**, 558–561 (2005)
31. De Sablet, T., Bertin, Y., Vareille, M., Girardeau, J.-P., Garrivier, A., Gobert, A.P., Martin, C.: Differential expression of stx2 variants in Shiga toxin-producing *Escherichia coli* belonging to seropathotypes A and C. *Microbiology* **154**, 176–86 (2008)
32. Fagerquist, C.K., Zaragoza, W.J.: Two Shiga toxin 2 subtypes in a single Shiga toxin-producing *Escherichia coli* analyzed by RT-qPCR, MALDI-TOF-TOF-MS, and top-down proteomic analysis. ASMS201416869.3977VER.1, Tuesday, June 17, 2014. Proceedings of the 62nd ASMS Conference on Mass Spectrometry and Allied Topics, Baltimore, MD, June 15–19 (2014)
33. Cooley, M.B., Jay-Russell, M., Atwill, E.R., Carychao, D., Nguyen, K., Quiñones, B., Patel, R., Walker, S., Swimley, M., Pierre-Jerome, E., Gordus, A.G., Mandrell, R.E.: Development of a robust method for isolation of shiga toxin-positive *Escherichia coli* (STEC) from fecal, plant, soil and water samples from a leafy greens production region in California. *PLoS One* **8**, e65716 (2013)
34. Rivas, E., Eddy, S.R.: A dynamic programming algorithm for RNA structure prediction including pseudoknots. *J. Mol. Biol.* **285**, 2053–2068 (1999)
35. 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)
36. Spengler, B.: Post-source decay analysis in matrix-assisted laser desorption / ionization mass spectrometry of biomolecules. *J. Mass Spectrom.* **32**, 1019–1036 (1997)
37. 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)
38. Demirev, P.A., Feldman, A.B., Kowalski, P., Lin, J.S.: Top-down proteomics for rapid identification of intact microorganisms. *Anal. Chem.* **77**, 7455–7461 (2005)
39. Boni, I.V., Isaeva, D.M., Musychenko, M.L., Tzareva, N.V.: Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. *Nucleic Acids Res.* **19**, 155–162 (1991)
40. Bear, D.G., Ng, R., Van Derveer, D., Johnson, N.P., Thomas, G., Schleich, T., Noller, H.F.: Alteration of polynucleotide secondary structure by ribosomal protein S1. *Proc. Natl. Acad. Sci. U.S.A.* **73**, 1824–1828 (1976)
41. Schlax, P.J., Worhunsky, D.J.: Translational repression mechanisms in prokaryotes. *Mol. Microbiol.* **48**, 1157–1169 (2003)
42. Deana, A., Belasco, J.G.: Lost in translation: the influence of ribosomes on bacterial mRNA decay. *Genes Dev.* **19**, 2526–2533 (2005)
43. Rauhut, R., Klug, G.: mRNA degradation in bacteria. *FEMS Microbiol. Rev.* **23**, 353–370 (1999)
44. Catherman, A.D., Skinner, O.S., Kelleher, N.L.: Top Down proteomics: facts and perspectives. *Biochem. Biophys. Res. Commun.* **445**, 683–693 (2014)
45. Kitova, E.N., Daneshfar, R., Marcato, P., Mulvey, G.L., Armstrong, G., Klassen, J.S.: Stability of the homopentameric B subunits of Shiga toxins 1 and 2 in solution and the gas phase as revealed by nanoelectrospray Fourier transform ion cyclotron resonance mass spectrometry. *J. Am. Soc. Mass Spectrom.* **16**, 1957–1968 (2005)
46. Zhang, Y., Liu, L., Daneshfar, R., Kitova, E.N., Li, C., Jia, F., Cairo, C.W., Klassen, J.S.: Protein-glycosphingolipid interactions revealed using catch-and-release mass spectrometry. *Anal. Chem.* **84**, 7618–7621 (2012)