JCM
Journals.ASM.org

Sequence-Level and Dual-Phase Identification of *Salmonella* Flagellum Antigens by Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS)

Keding Cheng,[a,b] Angela Sloan,[a] Julie Meakin,[a] Stuart McCorrister,[a] Morganne Jerome,[a] Garrett Westmacott,[a] Mike Drebot,[a] Celine Nadon,[a] J. David Knox,[a] Gehua Wang[a]

National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, Manitoba, Canada[a]; Department of Human Anatomy and Cell Sciences, Faculty of Medicine, University of Manitoba, Winnipeg, Manitoba, Canada[b]

Forty-three reference strains involving the 24 most common serovars of *Salmonella enterica* were examined by using a mass spectrometry-based H antigen typing platform (MS-H). The results indicate that MS-H can be used as a sensitive, rapid, and straightforward approach for the typing of *Salmonella* flagella at the molecular level without antiserum and phase inversion.

Salmonella bacteria are common food contaminants that can be lethal when consumed by humans. Typing of *Salmonella* is very important for identifying contaminated food sources and for surveillance of salmonellosis, a common human infection (1). The *Salmonella* genus comprises two species, with *S. enterica* occupying 99% of the detected isolates (2). Currently, there are 2,579 serovars according to a 2007 WHO *Salmonella* collaboration center report based on different formulae representing two main surface antigens, lipopolysaccharide (O antigens) and flagella (H antigens) (2). Interestingly, 80% of all *S. enterica* strains can be represented by only 10 different serovars (1, 3–6).

Serotyping of *Salmonella* is currently the gold standard and international language of *Salmonella* surveillance worldwide (2, 7). Although serotyping of these bacteria is a fairly simple test to run and observe, the preparation and procedures involved therein are time-consuming and laborious as they involve motility induction, phase suppression/inversion, and multistep agglutination reactions (7). No international standards pertain to antisera, and some are difficult to obtain, especially for those involved in the testing of rare and emerging strains of *Salmonella* (7). The serotyping procedure is more complicated for H typing than it is for O antigens because of frequent diphasic flagellum production (3). In such cases, a procedure called "phase inversion" must be applied whereby the production of one type of flagella is suppressed with antiserum while the other is identified. Motility induction is often employed to maximize flagellum production as well. For each phase of flagella, multiple factors need to be considered in order to determine clear results for closely related antigen complexes (2, 7, 8), with each factor requiring an agglutination reaction. For these reasons, our ISO-certified serotyping procedure normally takes 2 to 12 days to complete, depending on cell motility and the number of agglutination steps required.

Different molecular typing methods have been used in attempts to improve the speed, throughput, and quality of *Salmonella* typing, especially on the basis of flagellar genes. Among the most popular approaches are restriction fragment length polymorphism analysis (9–12), multiplex PCR (13, 14), and DNA microarray (15, 16). These approaches are promising in terms of speed and throughput but do not mirror the phenotypic properties and data quality of serotyping. Antibody array has also been investigated to improve the speed of *Salmonella* serotyping for common serovars (17), but this antibody-based approach still faces challenges when wide ranges of serovars, especially emerging serovars, are being observed. Recently, multilocus enzyme electrophoresis and multilocus sequence typing have been explored to link serotypes with housekeeping gene patterns (18–20), but neither method agreed well with results obtained by the traditional serotyping method. The bacterial identification method using whole-genome restriction patterns, pulsed-field gel electrophoresis, has been applied to some common serovars of *Salmonella* but could not achieve the quality of serotyping data because of variable patterns, even among strains with the same serovars (21). Lastly, MS techniques, especially matrix-assisted laser desorption ionization–time of flight MS, have been used to type *Salmonella* in recent years because of their speed and ease of use (22, 23). This fingerprinting-based method is also less than optimal because of the lack of consistent results and the lack of a complete fingerprinting database detailing the wide range of *Salmonella* serovars. Moreover, this platform cannot reach subspecies level analyses as can traditional serotyping (24).

Here, we explored a new mass spectrometry-based H antigen typing platform (MS-H) that uses liquid chromatography-tandem mass spectrometry (LC-MS/MS) to detect and type *Salmonella* flagella. For a detailed description of the method, see Text S1 in the supplemental material. Briefly, MS-H is based on our recent report of the *Escherichia coli* MS-H typing method (25). Flagellar samples were prepared by using overnight plate cultures of *Salmonella* reference strains without motility induction and phase inver-

Received 27 January 2014 **Returned for modification** 19 February 2014
Accepted 24 March 2014
Published ahead of print 2 April 2014
Editor: G. V. Doern
Address correspondence to Keding Cheng, keding.cheng@phac-aspc.gc.ca, or Gehua Wang, gehua.wang@phac-aspc.gc.ca.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/JCM.00242-14.

doi:10.1128/JCM.00242-14
The authors have paid a fee to allow immediate free access to this article.

TABLE 1 Data summary of MS-H-based flagellar typing of monophasic *Salmonella* reference strains

S. enterica serovar	Antigens	No. of strains	Top hit[a] (top hit/no. of tests)[b]
Enteritidis	9,12:g,m:−	3	g,m (6/6)
4,5,12:i:−	4,5,12:i:−	1	i (2/2)
Derby	4,12:f,g:−	1	f,g (2/2)
Monophasic Paratyphi B var. Java	4,12:b:−	2	b (4/4)
Paratyphi A	2,12:a:−	2	a (4/4)
Monophasic 4,[5],12,[27]:−:1,2	4,5,12:−:1,2	1	1,2 (2/2)
Alachua	35:z4,z23:−	1	z4,z23 (2/2)
Tennessee	6,7:z29:−	1	z29 (2/2)
Oranienburg	6,7,14:m,t:−	4	m,t (6/8); m,p,t,u (2/8)
Typhi	9,12:d:−	1	d (2/2)

[a] The top hit is the identified protein showing the highest Mascot score (signifying identification confidence) and emPAI value (signifying protein coverage, equal to 10 to the power of [the number of observed peptides divided by the number of observable peptides] minus one [reference 26]) and is regarded as the correct flagellar type.
[b] Top hit/no. of tests is how often the top hit was identified as such during MS-H testing.

sion. *Salmonella* flagella were trapped and purified on a 0.22-μm syringe filter membrane and then subjected to on-filter trypsin digestion and online LC-MS/MS detection of flagellin tryptic peptides. Peptide information was then assembled to obtain protein sequences that were compared to a curated database containing clear annotations of each flagellar type available for database setup. Twenty-four serovars from 43 strains occupying >80% of the most common serovars recently observed in our reference laboratory were selected for MS-H. Among the strains were 25 diphasic, 1 triphasic, and 17 monophasic isolates. Each strain was cultured twice to confirm the reproducibility and quality of the data output. Preparation of flagella was found to be more straightforward with *Salmonella* than with *E. coli*, as *Salmonella* strains are often motile with more stable yields of flagella than with *E. coli*, which can become nonmotile after long-term storage (25). The high-throughput platform operated continuously, with MS detection and database searches leading to flagellar identification for the first strain in only a few hours after sample preparation from culture growth.

We created a curated *Salmonella* flagellum database containing 385 entries by using those available in the literature and the NCBI nr database to aid in the database search. Since different serovars may share antigen/antigen factor complexes (2) and diphasic flagellum production is very common, *Salmonella* flagellar identification could be performed only at the antigen/antigen factor complex level by measuring sequence coverage and the associated identification confidence score. Both monophasic and diphasic isolates were identified with high confidence and accuracy because of the unique enrichment method used to extract high-purity flagella (25). As shown in Table 1 (also see Fig. S1 in the supplemental material) for monophasic *Salmonella* strains, all monophasic flagella were correctly identified with excellent reproducibility. The occasional identification of an "m,t" antigen complex as "m,p,t,u" was due to close sequence similarity, a scenario that would be difficult to differentiate during conventional serotyping as well. Successful and efficient identification of monophasic flagella should be very useful for rapid, high-quality H typing of *Salmonella enterica* serovars Enteritidis, Oranienburg, and Paratyphi, which are among the 10 most commonly reported serovars (1).

Identification of diphasic *Salmonella* flagellar types by MS-H was also shown to be faster and easier than by conventional serotyping. For an example of data output for flagellar identification of *S. enterica* serovar Newport (6,8:e,h:1,2), a common serovar with diphasic flagellum production, see Fig. S2 in the supplemental material. Here, four H antigen factors were identified concurrently in a single sample run with clear phylogenic differentiation, high confidence scores, peptide molecular weights, total and specific peptides identified, and relative quantities expressed as exponentially modified protein abundance indexes (emPAI; 10 to the power of [the number of observed peptides over the number of observable peptides] minus one) (26). Table 2 shows that all diphasic flagellar antigens were determined at the antigen factor complex/cluster level (8), with the majority identified at the individual-antigen level, particularly phase 1 H antigens. Complications arising during the characterization of different phase 2 *Salmonella* antigen factor 1 complexes (1,2; 1,5; 1,6; and 1,2,7) and phase 1 antigen groups ("r," "i," and "r,i") were due to extremely close sequence similarities (8), a phenomenon that often produces cross-reactions and hence ambiguous results during serotyping. As only 1/120 of the flagellin digest derived from one loopful of culture was used for routine sample loading during LC-MS/MS, an increase in data quality and accuracy would likely be achieved simply by repeating sample runs with a larger amount of protein digest (25). For the relationship between the loading amount and data accuracy when testing serovars Newport (6,8:e,h:1,2) and Infantis (6,7:r:1,5), respectively, see Tables S1 and S2 in the supplemental material. There was an interesting case in Table 2 in which MS-H on one isolate of Typhimurium kept getting serotyping factor "i" as "b" with high sequence coverage of 82% at a larger amount of sample loading. Although they are in the same α antigen cluster (8), this phenomenon needs to be explored further.

Phase 3 antigen z49 of serovar Infantis (6,7:r:1,5:z49), the only serovar containing phase 3 flagella, was likely not identified because of the lack of an available z49 sequence for comparison (Table 2). In addition, phase 2 flagella were not detected for some isolates, possibly because they were "spare flagella" and not consistently expressed (2, 7, 27). Further research on these observations is also necessary.

Overall, MS-H was found to be 100% accurate at the antigen level for 17 monophasic *Salmonella* strains and 100% accurate at the antigen cluster/complex level for 25 diphasic strains. Of these 25 diphasic strains, there was 75% accuracy for phase 1 antigens and 69% accuracy for unstable phase 2 antigens.

This report indicates that the fast, easy, sensitive, accurate, and high-throughput MS-H system can be applied to *Salmonella* bacterial H typing. This platform closely reflects traditional serotyping at the flagellar protein level with similar diagnostic sensitivity, excellent analytical sensitivity and specificity, and a more straightforward procedure. In addition, the observations from this report indicate that MS-H will be useful in *Salmonella* outbreak situations for the rapid screening of flagellar types. Lastly, as more flagellar genes are sequenced, particularly rare *Salmonella* serovars, this platform should decrease the labor and cost associated with antisera production and their maintenance in traditional serotyping.

TABLE 2 Data summary of MS-H-based flagellar typing of multiphasic *Salmonella* reference strains

S. enterica serovar	Antigens	No. of strains	H antigen top hit (top hit/no. of tests)[a]		Note
			Phase 1	Phase 2	
Typhimurium	4,5,12:i:1,2	2	r,i (2/4) b (2/4)	1,2 (2/4) 1,5 (2/4)	
Heidelberg	4,12:r:1,2	3	r (3/6) r,i (2/6) i (1/6)	1,2 (2/6) 1,5 (4/6)	
Typhi	9,12:d:z66	1	d (2/2)	z66 (2/2)	
Infantis	6,7:r:1,5:z49	1	r (2/2)	1,5 (2/2)	z49 sequence unavailable
Infantis	6,7:r:1,5	2	r (3/4) r,i (1/4)	1,5 (2/4) 1,2 (1/4) 1,(5),7 (1/4)	
Saintpaul	4,12:e,h:1,2	3	e,h (5/5)	1,2 (3/5) 1,5 (1/5) 1,6 (1/5)	
Hadar	6,8:z10:e,n,x	3	z10 (5/5)	e,n,x (4/5)	Once monophasic
Newport	6,8:e,h:1,2	2	e,h (4/4)	1,2 (4/4)	
Thompson	6,7:k:1,5	3	k (6/6)	1,5 (4/6) 1,2,7 (1/6)	Once monophasic
Javiana	9,12:l,z28:1,5	1	l,w (1/2) l,v (1/2)	1,5 (2/2)	
Uganda	3,15:l,z13:1,5	1	l, z13 (2/2)	enx15 (1/2)	Once monophasic
Schwarzengrund	4,12,27:d:1,7	1	d (2/2)	1,5 (1/2) 1,7 (1/2)	
Panama	9,12:l,v:1,5	1	l,v (2/2)	1,5 (2/2)	
Kentucky	8,20:i:z6	2	r,i (4/4)	z6 (4/4)	

[a] Top hit/no. of tests is how often the top hit was identified as such during MS-H testing. The top hit is the identified protein showing the highest Mascot score (signifying identification confidence) and emPAI value (signifying protein coverage, equal to 10 to the power of [the number of observed peptides divided by the number of observable peptides] minus 1 [reference 26]) and is regarded as the correct flagellar type.

ACKNOWLEDGMENTS

K.C. designed the method, performed flagellar extraction and sample preparation for MS-H, summarized data, and wrote the manuscript; A.S. was involved with critical writing of the paper; J.M. and M.J. performed reference strain selection and cell culture based on previous results obtained by ISO-accredited serotyping; S.M. executed mass spectrometry runs; M.D., C.N., and J.D.K. contributed project ideas; G. Westmacott contributed to mass spectrometry runs; and G. Wang contributed project ideas, method design, flagellar extraction, data summary, and manuscript writing.

REFERENCES

1. **Doyle ME, Kaspar C, Archer J, Klos R.** 2009. FRI briefings: white paper on human illness caused by *Salmonella* from all food and non-food vectors. Food Research Institute, University of Wisconsin—Madison, Madison, WI. http://fri.wisc.edu/docs/pdf/FRI_Brief_Salmonella_Human_Illness_6_09.pdf.
2. **Grimont PAD, Weill F-X.** 2007. Antigenic formulae of the *Salmonella* serovars, 9th edition. WHO Collaborating Centre for Reference and Research on *Salmonella*, Institut Pasteur, Paris, France. http://www.pasteur.fr/ip/portal/action/WebdriveActionEvent/oid/01s-000036-089.
3. **Anonymous.** 2013. Technical report: third external quality assurance scheme for *Salmonella* typing. European Food- and Waterborne Diseases and Zoonoses Network. European Centre for Disease Prevention and Control, Stockholm, Sweden. http://www.ecdc.europa.eu/en/publications/Publications/1204-TER-EQA-Salmonella-typing.pdf.
4. **Galanis E, Lo Fo Wong DM, Patrick ME, Binsztein N, Cieslik A, Chalermchikit T, Aidara-Kane A, Ellis A, Angulo FJ, Wegener HC, World Health Organization Global Salm-Surv.** 2006. Web-based surveillance and global *Salmonella* distribution, 2000-2002. Emerg. Infect. Dis. **12:**381–388. http://dx.doi.org/10.3201/eid1205.050854.
5. **Hendriksen RS, Vieira AR, Karlsmose S, Lo Fo Wong DM, Jensen AB, Wegener HC, Aarestrup FM.** 2011. Global monitoring of *Salmonella* serovar distribution from the World Health Organization Global Foodborne Infections Network Country Data Bank: results of quality assured laboratories from 2001 to 2007. Foodborne Pathog. Dis. **8:**887–900. http://dx.doi.org/10.1089/fpd.2010.0787.
6. **CDC.** 2011. National Enteric Disease Surveillance: national *Salmonella* surveillance overview. U.S. Department of Health and Human Services, CDC, Atlanta, GA. http://www.cdc.gov/nationalsurveillance/PDFs/NationalSalmSurveillOverview_508.pdf.
7. **Wattiau P, Boland C, Bertrand S.** 2011. Methodologies for *Salmonella*

enterica subsp. *enterica* subtyping: gold standards and alternatives. Appl. Environ. Microbiol. **77:**7877–7885. http://dx.doi.org/10.1128/AEM.05527-11.

8. **McQuiston JR, Parrenas R, Ortiz-Rivera M, Gheesling L, Brenner F, Fields PI.** 2004. Sequencing and comparative analysis of flagellin genes *fliC*, *fljB*, and *flpA* from *Salmonella*. J. Clin. Microbiol. **42:**1923–1932. http://dx.doi.org/10.1128/JCM.42.5.1923-1932.2004.
9. **Kilger G, Grimont PA.** 1993. Differentiation of *Salmonella* phase 1 flagellar antigen types by restriction of the amplified *fliC* gene. J. Clin. Microbiol. **31:**1108–1110.
10. **Dauga C, Zabrovskaia A, Grimont PA.** 1998. Restriction fragment length polymorphism analysis of some flagellin genes of *Salmonella enterica*. J. Clin. Microbiol. **36:**2835–2843.
11. **Dilmaghani M, Ahmadi M, Zahraei Salehi T, Talebi A, Darvishzadeh R.** 2010. PCR-restriction fragment length polymorphism analysis of *fljB* gene in *Salmonella enterica* subspecies *enterica* serovar Typhimurium isolated from avians. Iran. J. Microbiol. **2:**178–184.
12. **Paiva JB, Cavallini JS, Silva MD, Almeida MA, Ângela HL, Berchieri Junior A.** 2009. Molecular differentiation of *Salmonella* Gallinarum and *Salmonella* Pullorum by RFLP of *fliC* gene from Brazilian isolates. Rev. Bras. Cienc. Avicola **11:**271–276. http://dx.doi.org/10.1590/S1516-635X2009000400009.
13. **Herrera-León S, McQuiston JR, Usera MA, Fields PI, Garaizar J, Echeita MA.** 2004. Multiplex PCR for distinguishing the most common phase-1 flagellar antigens of *Salmonella* spp. J. Clin. Microbiol. **42:**2581–2586. http://dx.doi.org/10.1128/JCM.42.6.2581-2586.2004.
14. **Rajtak U, Leonard N, Bolton D, Fanning S.** 2011. A real-time multiplex SYBR green I polymerase chain reaction assay for rapid screening of *Salmonella* serotypes prevalent in the European Union. Foodborne Pathog. Dis. **8:**769–780. http://dx.doi.org/10.1089/fpd.2010.0768.
15. **McQuiston JR, Waters RJ, Dinsmore BA, Mikoleit ML, Fields PI.** 2011. Molecular determination of H antigens of *Salmonella* by use of a microsphere-based liquid array. J. Clin. Microbiol. **49:**565–573. http://dx.doi.org/10.1128/JCM.01323-10.
16. **Yoshida C, Franklin K, Konczy P, McQuiston JR, Fields PI, Nash JH, Taboada EN, Rahn K.** 2007. Methodologies towards the development of an oligonucleotide microarray for determination of *Salmonella* serotypes. J. Microbiol. Methods **70:**261–271. http://dx.doi.org/10.1016/j.mimet.2007.04.018.
17. **Cai HY, Lu L, Muckle CA, Prescott JF, Chen S.** 2005. Development of a novel protein microarray method for serotyping *Salmonella enterica* strains. J. Clin. Microbiol. **43:**3427–3430. http://dx.doi.org/10.1128/JCM.43.7.3427-3430.2005.
18. **McQuiston JR, Herrera-León S, Wertheim BC, Doyle J, Fields PI, Tauxe RV, Logsdon JM, Jr.** 2008. Molecular phylogeny of the salmonellae: relationships among *Salmonella* species and subspecies determined from four housekeeping genes and evidence of lateral gene transfer events. J. Bacteriol. **190:**7060–7067. http://dx.doi.org/10.1128/JB.01552-07.
19. **Sukhnanand S, Alcaine S, Warnick LD, Su WL, Hof J, Craver MP, McDonough P, Boor KJ, Wiedmann M.** 2005. DNA sequence-based subtyping and evolutionary analysis of selected *Salmonella enterica* serotypes. J. Clin. Microbiol. **43:**3688–3698. http://dx.doi.org/10.1128/JCM.43.8.3688-3698.2005.
20. **Achtman M, Wain J, Weill FX, Nair S, Zhou Z, Sangal V, Krauland MG, Hale JL, Harbottle H, Uesbeck A, Dougan G, Harrison LH, Brisse S, S. Enterica MLST Study Group.** 2012. Multilocus sequence typing as a replacement for serotyping in *Salmonella enterica*. PLoS Pathog. **8:**e1002776. http://dx.doi.org/10.1371/journal.ppat.1002776.
21. **Fendri I, Ben Hassena A, Grosset N, Barkallah M, Khannous L, Chuat V, Gautier M, Gdoura R.** 2013. Genetic diversity of food-isolated *Salmonella* strains through pulsed field gel electrophoresis (PFGE) and enterobacterial repetitive intergenic consensus (ERIC-PCR). PLoS One **8:**e81315. http://dx.doi.org/10.1371/journal.pone.0081315.
22. **Martiny D, Busson L, Wybo I, El Haj RA, Dediste A, Vandenberg O.** 2012. Comparison of the Microflex LT and Vitek MS systems for routine identification of bacteria by matrix-assisted laser desorption ionization–time of flight mass spectrometry. J. Clin. Microbiol. **50:**1313–1325. http://dx.doi.org/10.1128/JCM.05971-11.
23. **Kuhns M, Zautner AE, Rabsch W, Zimmermann O, Weig M, Bader O, Gross U.** 2012. Rapid discrimination of *Salmonella enterica* serovar Typhi from other serovars by MALDI-TOF mass spectrometry. PLoS One **7:**e40004. http://dx.doi.org/10.1371/journal.pone.0040004.
24. **Welker M, Moore ER.** 2011. Applications of whole-cell matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry in systematic microbiology. Syst. Appl. Microbiol. **34:**2–11. http://dx.doi.org/10.1016/j.syapm.2010.11.013.
25. **Cheng K, Drebot M, McCrea J, Peterson L, Lee D, McCorrister S, Nickel R, Gerbasi A, Sloan A, Janella D, Van Domselaar G, Beniac D, Booth T, Chui L, Tabor H, Westmacott G, Gilmour M, Wang G.** 2013. MS-H: a novel proteomic approach to isolate and type the *E. coli* H antigen using membrane filtration and liquid chromatography-tandem mass spectrometry (LC-MS/MS). PLoS One **8:**e57339. http://dx.doi.org/10.1371/journal.pone.0057339.
26. **Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, Mann M.** 2005. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol. Cell. Proteomics **4:**1265–1272. http://dx.doi.org/10.1074/mcp.M500061-MCP200.
27. **McQuiston JR, Fields PI, Tauxe RV, Logsdon JM, Jr.** 2008. Do *Salmonella* carry spare tyres? Trends Microbiol. **16:**142–148. http://dx.doi.org/10.1016/j.tim.2008.01.009.