Molekularbiologische und physiologische Untersuchungen zur Bedeutung phagenkodierter Sialinsäureesterasen von enterohämorrhagischen *Escherichia coli* (EHEC)

Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. *rer. nat.*)

Fakultät Naturwissenschaften
Universität Hohenheim

Institut für Lebensmittelwissenschaft und Biotechnologie
Fachgebiet Lebensmittelmikrobiologie und –hygiene

vorgelegt von
Nadja Saile (geb. Schairer)
aus Herrenberg

2019

Dekan:	Prof. Dr. Uwe Beifuß
1. berichtende Person:	Prof. Dr. Herbert Schmidt
2. berichtende Person:	Prof. Dr. Lutz Fischer

Eingereicht am: 24.01.2019

Mündliche Prüfung am: 23.05.2019

Die vorliegende Arbeit wurde am 25.04.2019 von der Fakultät Naturwissenschaften der Universität Hohenheim als „Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften“ angenommen.

Eidesstattliche Versicherung gemäß § 7 Absatz 7 der Promotionsordnung der Universität Hohenheim zum Dr. rer. nat.

1. Bei der eingereichten Dissertation zum Thema „Molekularbiologische und physiologische Untersuchungen zur Bedeutung phagenkodierter Sialinsäureesterasen von enterohämorrhagischen *Escherichia coli* (EHEC)“ handelt es sich um meine eigenständig erbrachte Leistung.

2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.

3. Ich habe nicht die Hilfe einer kommerziellen Promotionsvermittlung oder -beratung in Anspruch genommen.

4. Die Bedeutung der eidesstattlichen Versicherung und der strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt.

Die Richtigkeit der vorstehenden Erklärung bestätige ich: Ich versichere an Eides Statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

_____ _____

Ort und Datum Unterschrift

Danksagung

Zuerst möchte ich mich bei Herrn Prof. Dr. Herbert Schmidt bedanken für die Überlassung des Themas, die Möglichkeit der Bearbeitung an seinem Fachgebiet, die Unterstützung und die jahrelange gute Zusammenarbeit.

Herrn Prof. Dr. Lutz Fischer danke ich für die freundliche Übernahme des Zweitgutachtens.

Ein besonderer Dank geht an Frau Dipl.-Ing. Dr. Agnes Weiß für fachliche Diskussionen und Anregungen sowie ihre stets guten Ratschläge.

Ich bedanke mich bei meinen Kolleginnen und Kollegen vom Fachgebiet Lebensmittelmikrobiologie und -hygiene für die schöne Zeit, die witzigen Momente im Laboralltag und die Motivation. Insbesondere danke ich den technischen Assistenten Heike Popovitsch, Melanie Schneider, Claudia Lis und Markus Kranz, die mich bei meiner Arbeit im Labor unterstützten.

Herzlichen Dank an alle Master- und Bachelorstudierenden, die an der Thematik motiviert mitarbeiteten und in diesem Rahmen ihre Abschlussarbeiten anfertigten und an alle Co-Autoren für die gelungene Zusammenarbeit bei der Erstellung der Publikationen.

Danke an die Deutsche Forschungsgemeinschaft (Schm1360/6-1) für die Bereitstellung der finanziellen Mittel, die zum Gelingen der Arbeit nötig waren.

Am Ende danke ich meinem Mann Uwe, meiner Schwester Rebekka und dem Rest meiner Familie, dass sie manchmal mit mir in die Welt der EHEC eintauchten und für deren Hilfe, Rückhalt und Geduld.

Inhaltsverzeichnis

Zusammenfassung

Enterohämorrhagische *Escherichia coli* (EHEC) können im Menschen Komplikationen wie hämorrhagische Colitis oder das lebensbedrohliche hämolytisch urämische Syndrom auslösen. Ihre Hauptvirulenzfaktoren sind Shiga Toxine (Stx) und ein Typ-III-Sekretionsystem. Zur Kolonisation des Dickdarms stehen EHEC mit der Darmmikrobiota im Wettbewerb um limitierte Energiequellen wie die Sialinsäure 5-*N*-Acetylneuraminsäure (Neu5Ac). Für den Neu5Ac-Katabolismus besitzen *E. coli* die Gene *nanA*, *nanT*, *nanE*, *nanK*, *nagA* und *nagB*. Sialinsäuren kommen endständig an Muzinen des Mukus im Dickdarm vor. *E. coli* bilden keine Sialidasen, die zur Freisetzung von Neu5Ac aus Muzin nötig sind. Jedoch kann gebundenes Neu5Ac durch Sialidasen von anaeroben Bakterien im Darm abgespalten werden und sind dann verfügbar. Viele *E. coli* Stämme exprimieren das *nanCMS* Operon. NanC ist ein Porinprotein, NanM ist eine Mutarotase und NanS ist eine *O*-Acetyl Esterase. NanS konvertiert 5-*N*-Acetyl-9-*O*-Acetylneuraminsäure (Neu5,9Ac_2) und Neu5,8Ac_2 zu Neu5Ac und ist maßgeblich an der Energiegewinnung aus diesen Substraten beteiligt. In Prophagen von EHEC und anderer pathogener *E. coli* kommen multiple *nanS*-homologe offene Leserahmen vor, die in dieser Arbeit als *nanS*-p (p = Phage) deklariert wurden. In dieser Dissertation sollte die Frage geklärt werden, warum EHEC multiple *nanS*-p Gene besitzen und ob die entsprechenden Proteine zur Energiegewinnung und zu einem Wettbewerbsvorteil für EHEC Bakterien beitragen. Außerdem sollte die Beteiligung der Sialidase BTSA von *Bacteroides thetaiotaomicron* an der NanS-p abhängigen Verwertung von Muzin untersucht werden.
Für die Analysen wurde der pathogene *E. coli* O157:H7 Stamm EDL933 sowie die pathogenen O104:H4 Stämme LB226692 und C227-11φcu (Variante von C227-11, die keinen *stx*-Prophagen besitzt) und die apathogenen Stämme AMC 198 (*nanS*$^+$, *nanS*-p$^-$) und C600Δ*nanS* (*nanS*$^-$, *nanS*-p$^-$) verwendet. Die Chromosomensequenzen von EDL933 und LB226692 wurden *in silico* analysiert. NanS-p2 und NanS-p4 aus EDL933 wurden rekombinant hergestellt und das pH- und das Temperaturoptimum, sowie potenzielle Substrate analysiert. *nanS*/*nanS*-p Deletionsmutanten von EDL933 und C227-11φcu wurden konstruiert und mit Neu5,9Ac_2 oder Muzin kultiviert. Dabei wurden die Verläufe der Wachstumskurven durch Trübungsmessung

oder Lebendkeimzahlbestimmung analysiert. In Co-Kultivierungsversuchen wurden dem Medium der Stamm AMC 198 und/oder BTSA und/oder NanS-p zugesetzt.
Das Chromosom von EDL933 enthält das chromosomale *nanS* und sieben *nanS*-p Gene, während in LB226692 kein *nanS*, aber fünf *nanS*-p Gene vorkommen. Es konnte bestätigt werden, dass alle *nanS*-p Gene in der spät regulierten Region von Prophagen lokalisiert sind. Die putativen NanS-p Proteine tragen die Domänen 303, die für die Esterasefunktion in NanS bekannt ist, und die unbekannte Domäne 1737. Die untersuchten rekombinanten NanS-p Proteine de-*O*-acetylierten die Substrate Neu5,9Ac_2 und Rinderspeicheldrüsenmuzin und zeigten ein Temperatur- und pH-Optimum von 40-50 °C bzw. 7-9. Die Generationszeiten der Deletionsmutanten stiegen mit der Anzahl der deletierten *nanS*-p Gene im Chromosom und somit konnte ein Gen-Dosis-Effekt nachgewiesen werden. Durch extrazelluläre Supplementierung mit NanS-p konnte die ursprüngliche Wachstumskinetik des Wildtyps wiederhergestellt werden. Die einzelne Deletion von *nanS* im EDL933 Chromosom (EDL933Δ*nanS*) änderte die Wachstumskinetik des Stammes mit Neu5,9Ac_2 nicht. NanS ist für eine Vermehrung von EDL933 mit Neu5,9Ac_2 also nicht notwendig. C600Δ*nanS*, konnte Neu5,9Ac_2, wie erwartet, nicht katabolisieren. Den ursprünglichen Verlauf der Wachstumskurven einer C600 Kultivierung zeigte C600Δ*nanS* jedoch durch NanS-p Zusatz im Medium. In einem Co-Kultivierungsexperiment stieg die Lebendkeimzahl von C227-11φcu an, während sich AMC 198 kaum vermehrte. Erst nachdem alle *nanS*-p Gene im C227-11φcu Chromosom deletiert waren, stieg die Lebendkeimzahl von AMC 198. In Muzin-haltigem Medium mit BTSA als Zusatz, konnte sich EDL933 und EDL933Δ*nanS*, nicht aber EDL933Δ*nanS*Δ*nanS*-p1a-p7, vermehren. Dies zeigte, dass die Kombination aus BTSA und NanS-p das Wachstum von EDL933 in einer Muzin-haltigen Umgebung unterstützt.
Die Ergebnisse dieser Arbeit zeigten zum ersten Mal die wichtige Bedeutung der Prophagen-kodierten *O*-Acetyl Esterasen für EHEC Bakterien und sind eine hervorragende Basis für weitere Arbeiten, die zu einem tieferen wissenschaftlichen Verständnis des Sialinsäurekatabolismus in EHEC und anderen pathogenen *E. coli* beitragen können. Sie identifizieren Sialinsäuren als Substrate einer potenzielle Nährstoffnische im Darm und eröffnen ein mögliches Ziel für therapeutische Ansätze.

Abstract

Enterohemorrhagic *Escherichia coli* (EHEC) are responsible for severe disease in humans such as hemorrhagic colitis or the life-threatening hemolytic uremic syndrome. The main virulence factors are Shiga toxins (Stx) and a type-III-secretion system. For colonization of the colon, they have to compete with the intestinal microbiota for limiting substrates such as 5-*N*-acetyl neuraminic acid (Neu5Ac). For the catabolism of Neu5Ac, *E. coli* possess the genes *nanA*, *nanT*, *nanE*, *nanK*, *nagA,* and *nagB*. Many sialic acids are terminally bound to mucins of mucus in the colon. *E. coli* is not able to use bound Neu5Ac, because it cannot express sialidases for cleavage of Neu5Ac from mucin. However, sialic acids can be cleaved by sialidases from anaerobic bacteria in the colon and are then available. Many *E. coli* strains encode the *nanCMS* operon. NanC is a porin protein, NanM is a mutarotase and NanS is an *O*-acetyl esterase. NanS converts 5-*N*-acetyl-9-*O*-acetyl neuraminic acid (Neu5,9Ac_2) and Neu5,8Ac_2 to Neu5Ac and offers a substrate niche. There are multiple prophage-located *nanS*-homologous open reading frames in the chromosomes of EHEC, that are designated as *nanS*-p (p = phage) in this study. The question of this thesis was, why EHEC possess multiple *nanS*-p genes, if their corresponding proteins are used for energy production and for a competition advantage of EHEC. Furthermore, the influence of the sialidase BTSA of *Bacteroides thetaiotaomicron* on the NanS-p dependent utilization of mucin should be investigated.
The strains used in this study were the pathogenic *E. coli* O157:H7 strain EDL933 as well as the pathogenic O104:H4 strains LB226692 and C227-11φcu (variant of C227-11 cured in its *stx*-prophage), and the apathogenic *E. coli* strains AMC 198 ($nanS^+$, $nanS$-p^-) and C600Δ*nanS* ($nanS^-$, $nanS$-p^-). The chromosome sequences of EDL933 and LB226692 were analyzed *in silico*. NanS-p2 and NanS-p4 of EDL933 were expressed recombinantly and the pH- and temperature-optimum as well as potential substrates were investigated. *nanS*/*nanS*-p deletion mutants of EDL933 and C227-11φcu were generated and cultivated with Neu5,9Ac_2 or mucin. The courses of the growth curves were analyzed by turbidity measurement or determination of the viable cell counts. In co-cultivation experiments the medium was supplemented with the AMC 198 strain and/or BTSA and/or NanS-p.

EDL933 possess the chromosomal *nanS* and seven *nanS*-p genes, while LB226692 has no *nanS*, but five *nanS*-p genes. It was confirmed, that all *nanS*-p genes are located in the late regulated gene region of the prophages. The putative NanS-p proteins include the domain 303, which is known for its esterase activity in NanS, and the domain of unknown function 1737. The analyzed recombinant NanS-p proteins de-*O*-acetylated the substrates Neu5,9Ac_2 and mucin of bovine submaxillary gland and showed a temperature- and pH-optimum of 40-50 °C and 7-9, respectively. A gene-dose effect was identified because the generation times of the deletion mutant strains increased with the number of deleted *nanS*-p genes in the chromosomes. After extracellular supplementation with NanS-p, the mutant strains regained the original growth kinetic of the wildtype strain. The mono-deletion of *nanS* in EDL933 (EDL933Δ*nanS*) caused no chances in the growth kinetic of the strain with Neu5,9Ac_2. Therefore, NanS is not necessary for growth of EDL933 on Neu5,9Ac_2. C600Δ*nanS* could not catabolize Neu5,9Ac_2, as expected. The original course of the growth curve of a C600 cultivation was recuperated by C600Δ*nanS* if NanS-p was added to the medium. In a co-cultivation the viable cell count of C227-11φcu increased, while AMC 198 increased hardly. Only after deletion of all *nanS*-p genes in C227-11φcu, the viable cell count of AMC 198 increased. In a mucin-containing medium with BTSA as supplement EDL933 and EDL933Δ*nanS* could grow in contrast to EDL933Δ*nanS*Δ*nanS*-p1a-p7. This shows, that the combination of BTSA and NanS-p supports the growth of EDL933 in a mucin-containing environment.

The results of this work showed for the first time the importance of prophage-encoded *O*-acetyl esterases for EHEC bacteria and are an excellent basis for further work that contribute to a profound scientific understanding of the sialic acid catabolism in EHEC and other pathogenic *E. coli*. They identify sialic acids as substrates of a potential nutrient niche in the gut and establish a potential location for therapeutic approaches.

Publikationsliste

Die verschiedenen Teile dieser Arbeit wurden in Fachjournalen publiziert oder auf Fachkongressen vorgestellt. Die Erstautoren bzw. präsentierende Personen wurden nachfolgend durch Unterstreichung hervorgehoben.

Originalarbeiten in Peer-Review-Journalen

Saile, N., Voigt, A., Kessler, S., Stressler, T., Klumpp, J., Fischer, L., Schmidt, H., 2016. *Escherichia coli* O157:H7 strain EDL933 harbors multiple functional prophage-associated genes necessary for the utilization of 5-*N*-acetyl-9-*O*-acetyl neuraminic acid as a growth substrate. Appl. Environ. Microbiol. 82:5940-5950.

Saile, N., Schwarz, L., Eißenberger, K., Klumpp, J., Fricke, F.W., Schmidt, H., 2018. Growth advantage of *Escherichia coli* O104:H4 strains on 5-*N*-acetyl-9-*O*-acetyl neuraminic acid as a carbon source is dependent on heterogeneous phage-borne NanS-p esterases. Int. J. Med. Microbiol. 308:459-468.

Feuerbaum, S.*, Saile, N.*, Pohlentz, G., Müthing, J., Schmidt, H., 2018. De-*O*-acetylation of mucin-derived sialic acids by recombinant NanS-p esterases of *Escherichia coli* O157:H7 strain EDL933. Int. J. Med. Microbial. 308:1113-1120. *contributed equally to this work.

Vorträge auf Fachkongressen

Schairer, N., Schmidt, H., 2014. Effect of prophage genes on the growth of enterohemorrhagic *Escherichia coli* O157:H7 with 5-*N*-acetyl-9-*O*-acetylneuraminic acid as a carbon source. 4. Gemeinsame Konferenz der Deutschen Gesellschaft für Hygiene und Mikrobiologie e. V. und der Vereinigung für Allgemeine und Angewandte Mikrobiologie, Dresden, Deutschland.

Saile, N., Voigt, A., Fischer, L., Schmidt, H., 2015. Characterization of prophage-encoded esterases of *Escherichia coli* O157:H7 and their role in substrate utilization. 67. Jahrestagung der Deutschen Gesellschaft für Hygiene und Mikrobiologie e. V., Münster, Deutschland.

Saile, N., Nübling, S., Voigt, A., Stressler, T., Fischer, L., Schmidt, H., 2016. Prophagenkodierte Enzyme von *E. coli* O157:H7 Stamm EDL933 ermöglichen die Nutzung von 5-*N*-Acetyl-9-*O*-Acetyl-Neuraminsäure als Kohlenstoffquelle. V. EHEC-Workshop, Nördlingen, Deutschland.

Saile, N., Schmidt, H., 2016. Enterohämorrhagische *E. coli* O157:H7 Esterase fördert Wachstum eines apathogenen *E. coli* Stamms auf Mukussubstrat 5-*N*-Acetyl-9-*O*-Acetyl-Neuraminsäure. 16. Fachsymposium Lebensmittelmikrobiologie, Stuttgart-Hohenheim, Deutschland.

Saile, N., Voigt, A., Nübling, S., Kessler, S., Stressler, T., Klumpp, J., Fischer, L., Schmidt, H., 2017. *Escherichia coli* O157:H7 strain EDL933 expresses several *O*-acetyl esterases being involved in sialic acid catabolism. 5. Gemeinsame Konferenz der Deutschen Gesellschaft für Hygiene und Mikrobiologie e. V. und der Vereinigung für Allgemeine und Angewandte Mikrobiologie, Würzburg, Deutschland.

Posterpräsentationen auf Fachkongressen

Saile, N., Voigt, A., Fischer, L., Schmidt, H., 2016. Characterization of 5-*N*-acetyl-9-*O*-acetyl neuraminic acid-esterases from enterohemorrhagic *E. coli* O157:H7. Food-Micro 2016, Dublin, Irland.

Saile, N., Schwarz, L., Eißenberger, K., Klumpp, J., Fricke, F., Schmidt H., 2018. Competitive growth of *E. coli* O104:H4 strain with 5-*N*-acetyl-9-*O*-acetyl neuraminic acid as a carbon source. 70. Jahrestagung der Deutschen Gesellschaft für Hygiene und Mikrobiologie e. V., Bochum, Deutschland.

Feuerbaum, S., Saile, N., Pohlentz, G., Müthing, J., Schmidt, H., 2018. De-*O*-acetylation of sialic acids by recombinant NanS-p1, NanS-p2 and NanS-p4 esterases of *Escherichia coli* O157:H7 strain EDL933. 10th International Symposium on Shiga Toxin (Verocytotoxin) producing *Escherichia coli* infections, Florenz, Italien.

Schmidt, H., Saile, N., 2018. Analysis of multiple phage-borne NanS-proteins of enterohemorrhagic *E. coli* O157:H7 and O104:H4 strains. 10th International Symposium on Shiga Toxin (Verocytotoxin) producing *Escherichia coli* infections, Florenz, Italien.

Weitere wissenschaftliche Veröffentlichungen

Neben den oben genannten Veröffentlichungen wurden weitere wissenschaftliche Beiträge im Bearbeitungszeitraum der Dissertation veröffentlicht, die nicht Bestandteil der vorliegenden Dissertation sind.

Originalarbeiten in Peer-Review-Journalen

Nübling, S., Wohlt, D., Saile, N., Weiss, A., Schmidt, H., 2017. Antimicrobial effect of lauroyl arginate ethyl on *Escherichia coli* O157:H7 and *Listeria monocytogenes* on red oak leaf lettuce. Eur. Food Res. Technol. 243:879–887.

Saile, N.*, Schuh, E.*, Semmler, T., Eichhorn, I., Wieler, L.H., Bauwens, A., Schmidt, H., 2018. Determination of virulence and fitness genes associated with the *pheU*, *pheV* and *selC* integration sites of LEE-negative food-borne Shiga toxin-producing *Escherichia coli* strains. Gut Pathog. doi: 10.1186/s13099-018-0271-8. *contributed equally to this work

Posterpräsentationen auf Fachkongressen

Eißenberger, K., Drissner, D., Saile, N., Weiss, A., Schmidt, H., 2017. Enterohemorrhagic *Escherichia coli* O157:H7 strain Sakai can be internalized into the root tissues of corn salad and lettuce following contamination of the soil under greenhouse conditions. Microbiology and Infection 2017, Würzburg, Deutschland.

Angaben zum Eigenanteil der Publikationen

Im Rahmen dieser Arbeit wurden drei Manuskripte erstellt, die die Vorgaben für eine kumulative Dissertation erfüllen. Bei der Erstellung waren mehrere Co-Autoren beteiligt. **Herbert Schmidt** war der Antragsteller des DFG Projektes (Schm1360/6-1). Er war maßgeblich an der Konzeption der Versuche aller drei Manuskripte (Kapitel 2-4) beteiligt, unterstützte bei der Erstellung der Manuskripte und korrigierte sie.

Manuskript I (Kapitel 2)

Nadja Saile. Ich bereitete die genomische DNA für die Sequenzierung vor, stellte die *E. coli* O157:H7 Deletionsmutanten her, führte die Wachstumsversuche durch, unternahm die Mutagenese und Klonierung von *nanS*-p2, exprimierte NanS-p1, NanS-p2 und NanS-p4, reinigte diese auf und führte die Bradford Analysen durch. Ich führte die funktionelle und enzymatische Charakterisierung der Enzyme durch. Die bioinformatischen Analysen zur Sequenz NZ_CP008957 wurden von mir, nach Vorlage von Anja Voigt, durchgeführt. Ich schrieb und überarbeitete das Manuskript.

Anja Voigt führte bioinformatischen Analysen mit der Sequenz AE005174 durch, die als Vorlage für die bioinformatischen Analysen der Sequenz NZ_CP008957 des Manuskripts dienten. Sie klonierte *nanS*-p4 und überarbeitete das Manuskript.

Sarah Kessler stellte die Deletionsmutante *E. coli* C600Δ*nanS* her.

Timo Stressler beaufsichtige die enzymatische Charakterisierung und unterstützte bei der Auswertung der Daten. Er überarbeitete das Manuskript.

Jochen Klumpp führte die Sequenzierung zur Sequenz CP015855 durch und war an der Überarbeitung des Manuskripts beteiligt.

Lutz Fischer unterstützte bei der Erhebung der enzymatischen Daten und überarbeitete das Manuskript.

Manuskript II (Kapitel 3)

Nadja Saile. Ich bereitete die genomische DNA für die Genomsequenzierung vor, führte alle *in silico* Analysen durch, stellte die *E. coli* O104:H4 Deletionsmutanten her und führte die Kompetitionsversuche durch. Ich klonierte BT_0455, exprimierte

und reinigte BTSA und führte die Bradford Analyse für BTSA durch. Ich plante und beaufsichtige die Enzymcharakterisierung von NanS-p13 und wertete die Daten aus. Ich schrieb und überarbeitete das Manuskript.

Lisa Schwarz klonierte *nanS*-p13, exprimierte NanS-p13, reinigte es auf und führte die Bradford Analyse dazu durch.

Kristina Eißenberger konstruierte das Plasmid pKEC1.5.

Jochen Klumpp führte die Sequenzierungen zu den Sequenzen CP024992-CP024994 durch und war an der Überarbeitung des Manuskripts beteiligt.

Florian W. Fricke annotierte die Sequenzen CP024992-CP024994 und überarbeitete das Manuskript.

Manuskript III (Kapitel 4)

Bei der Erstautorenschaft zu Manuskript III handelt es sich um eine geteilte Erstautorenschaft zwischen Stefanie Feuerbaum und Nadja Saile.

Stefanie Feuerbaum exprimierte und reinigte NanS und führte die Bradford Analyse durch. Sie führte die enzymatischen Umsetzungen der Sialinsäuren und von Muzin durch NanS, NanS-p Proteinen und BTSA durch. Stefanie Feuerbaum führte die HPTLC Analysen durch und wertete die HPTLC und nanoESI Daten aus. Sie schrieb und überarbeitete das Manuskript für die genannten Versuche.

Nadja Saile. Ich führte die Kultivierungsversuche mit und ohne BTSA und NanS-p in Muzin durch und wertete die Daten aus. Ich schrieb und überarbeitete das Manuskript für die genannten Versuche.

Gottfried Pohlentz führte die nanoESI Analysen durch und unterstützte bei der Auswertung der nanoESI Daten.

Johannes Müthing überarbeitete das Manuskript und hatte die Leitung der nanoESI Analysen an der Universität Münster.

Ort und Datum

Unterschrift des Betreuers

1 Einleitung

1.1 Enterohämorrhagische *Escherichia coli* und andere Shiga Toxin-bildenden *E. coli*

Enterohämorrhagische *Escherichia coli* (EHEC) sind Krankheitserreger, die hauptsächlich durch kontaminierte Lebensmittel auf den Menschen übertragen werden. Die häufigsten Symptome einer EHEC Infektion sind wässriger Durchfall mit abdominalen krampfartigen Schmerzen. Seltener tritt eine hämorrhagische Colitis (HC) auf und bei schweren Verläufen entwickelt sich das hämolytisch-urämische Syndrom (HUS), einhergehend mit mikroangiopathischer hämolytischer Anämie, Thrombozytopenie und akutem Nierenversagen (Nataro & Kaper, 1998; Tarr, 2009). Besonders Kinder, ältere Personen und immunsupprimierte Menschen sind betroffen (Karmali, 1989). Die Infektionsdosis ist mit weniger als 100 Zellen gering (Tuttle et al., 1999). Als Kommensalen kommen EHEC vor allem in Rindern vor und gelangen über Fäzes in die Umwelt (Nataro & Kaper, 1998). Eine Gefahr für den Menschen besteht, wenn nicht durchgegartes kontaminiertes Fleisch oder kontaminierte roh verzehrbare Produkte, wie Nutzpflanzen, Rohmilchkäse (Zweifel et al., 2010; Costard et al., 2017) oder Mettwurst (Paton et al., 1996), konsumiert werden. Durch den direkten Kontakt zu Tieren (Chapman et al., 2000; Heuvelink et al., 2002) oder durch Mensch-zu-Mensch Übertragungen (Reida et al., 1994, Scavia et al., 2018) kam es in der Vergangenheit ebenfalls zu Erkrankungen.

Im klassischen Sinne sind EHEC Shiga Toxin (Stx)-bildende *E. coli* (STEC), die „Attaching und Effacing (A/E)“-Läsionen an Enterozyten auslösen, ein circa 60 MDa großes Plasmid besitzen und zu HC und HUS führen (Levine, 1987). Für den A/E Phänotyp ist der „Locus of Enterocyte Effacement“ verantwortlich (Jerse et al., 1990) und beim besagten Plasmid handelt es sich um das pO157 (Toth et al., 1990). „STEC“ dagegen umfasst alle *E. coli* Bakterien, die eine oder beide Gruppen der Stx-Gene (stx_1, stx_2) besitzen. Dabei können STEC humanpathogen sein, müssen dies aber nicht zwangsläufig (Nataro & Kaper, 1998).

Obwohl die Mehrheit der EHEC Infektionen sporadisch ist, kam es in der Vergangenheit weltweit zu einigen größere Ausbrüche. Der erste, über den berichtet wurde,

fand im Jahr 1982 in Michigan und Oregon, USA statt, bei dem die Erkrankten zuvor kontaminiertes Fleisch bei einer Fast-Food-Kette konsumiert hatten (Riley et al., 1983). Der verantwortliche Stamm, der während des Ausbruchs isoliert wurden gehörten zum Serotyp O157:H7 (O'Brien et al., 1993; Riley et al., 1983). In dieser Arbeit wurde das Isolat EDL933 verwendet, das aus Hackfleisch in Michigan isoliert wurde und mit diesem Ausbruch in Zusammenhang steht (Kapitel 2 und 4). *E. coli* O157:H7 sind die häufigsten Verursacher von HUS (Karch et al., 2005; Tarr, 2009). Weitere Ausbrüche mit Stämmen dieses Serotyps ereigneten sich im Jahr 1996 in Sakai in Japan durch den Verzehr von kontaminierte Rettichsprossen (Michino et al., 1999), 2006 in Nordamerika durch Verzehr von kontaminiertem Spinat (Grant et al., 2008) oder 2014 in Alberta in Kanada durch Verzehr von kontaminiertem Schweinfleisch (Honish et al., 2017). Neben O157:H7 kamen die Serotypen O26:H11, O45:H2, O103:H2, O111:H8, O121:H19 und O145:H28 häufig im Zusammenhang mit HUS und HC vor (Karch et al., 2005; Mora et al., 2007).

Im Jahr 2011 ereignete sich ein Ausbruch von Durchfallerkrankungen und HUS in Deutschland und anderen Teilen Europas (Frank et al., 2011; Anonym a, 2012; King et al., 2012). Infektionsquelle waren höchstwahrscheinlich mit *E. coli* kontaminierte Bockshornkleesprossen (Anonym a, 2012; King et al., 2012). Der während des Ausbruchs isolierte Stamm LB226692 des Serotyps O104:H4 (Mellmann et al., 2011) war kein EHEC im klassischen Sinn. Er konnte Stx2a bilden, besaß aber keinen Locus of Enterocyte Effacement und kein Plasmid pO157 (Bielaszewska et al., 2011; Brzuszkiewicz et al., 2011). Aufgrund seiner genetischen Ausstattung ist er den enteroaggregativen *E. coli* (EAEC) näher verwandt als den EHEC. Der Stamm bildet die aggregativen Adhärenzfimbrien vom Typ I (AAF/I), die in Zellkulturuntersuchungen mit HEp-2 Zellen für das EAEC typische Muster der „gestapelten Backsteine" verantwortlich sind (Bielaszewska et al., 2011; Tietze et al., 2015). Desweiteren bildet er einige Extended-Spectrum Betalaktamasen, die zur Multiresistenz des Stammes beitragen und die Serinproteaseautotransporter SepA, SigA und Pic, die an der Kolonisation und Zerstörung von Mukus beteiligt sind (Brzuszkiewicz et al., 2011; Mellmann et al., 2011; Rasko et al., 2011). Während des Ausbruchs erkrankten in Deutschland 3793 Menschen, 827 davon entwickelten ein HUS und 53 starben in Folge (Anonym a, 2012). Weltweit und durch Folgeinfektionen sind die

Zahlen höher. Im Rahmen dieser Arbeit wurde eine zusammenhängende Genomsequenz von LB226692 an der ETH Zürich erstellt. In der vorliegenden Studie wurden *in silico* Analysen mit dieser Genomsequenz unternommen (Kapitel 3). *In vitro* Untersuchungen erfolgten mit einem anderen Isolat des Ausbruchs (C227-11), bei dem der stx_{2a}-Prophage künstlich entfernt wurde (C227-11φcu) (Kapitel 3; Zangari et al., 2013).

Bei der Behandlung von STEC Infektionen wird die Gabe von Antibiotika vermieden, da Antibiotika die *stx*-Expression induzieren und die Entwicklung von HUS fördern (Grif et al., 1998; Köhler et al., 2000; Wong et al., 2000; Zhang et al. 2000; Herold et al., 2005; Freedman et al., 2016). Therapeutische Ansätze zielen darauf ab, STEC zu eliminieren ohne die *stx*-Expression zu induzieren, die Genexpression in STEC zu stören, die Generierung der Stx Rezeptoren zu limitieren oder Toxin-Bindung, Transport, Prozessierung oder Aktivität in der Zelle zu unterbinden (Melton-Celsa & O'Brien, 2014; Pacheco et al., 2018). Auch der Einsatz monoklonaler Antikörper gegen Stx wird diskutiert (Melton-Celsa & O'Brien, 2014).

1.2 Virulenzfaktoren von EHEC

EHEC und andere pathogene *E. coli* unterscheiden sich von kommensalen *E. coli* durch Bildung einer Reihe von Virulenzfaktoren. Ergänzend zum konservierten *E. coli* Kerngenom besitzen sie zusätzliche genetische Elemente. Der größte Teil ist horizontal erworbene fremde DNA wie genomische Inseln, Prophagen oder Transposons (Dobrindt, 2005). Zusätzliche extrachromosomale DNA, wie Plasmide, tragen weiter zur Diversität der Stämme bei (Dobrindt, 2005). Proteine, die von diesem flexiblen Genpool kodiert werden, beeinflussen die Fitness und Virulenz der Stämme und tragen zum Überleben bei. Die Adaptierung an ökologische Nischen und die Bestimmung des Infektionsorts sind von der genomischen Flexibilität abhängig (Brzuszkiewicz et al., 2009; Dobrindt et al., 2010).

Charakteristische Virulenzfaktoren der EHEC sind ihre Stx, die von Prophagen kodiert werden (siehe Kapitel 1.3) und das Typ-III-Sekretionssystem (T3SS), das Adhäsin Intimin und Effektorproteine, die alle von einer Pathogenitätsinsel (PAI) kodiert werden. Das pO157 kodiert gleich mehrere Virulenzfaktoren wie das Hämolysin EhxA (Schmidt et al., 1994), die Katalase-Peroxidase KatP (Brunder et al., 1996),

das Typ-II-Sekretionssystem Etp (Schmidt et al., 1997), die Serinprotease EspP (Brunder et al., 1997), das Adhäsin ToxB (Tatsuno et al., 2001), die Zinkmetalloprotease StcE (Lathem et al., 2002) und das *eae* konservierte Fragment Ecf (Yoon et al., 2005).

Neben diesen typischen Virulenzfaktoren exprimieren manche EHEC Stämme noch weitere Toxine wie das Subtilase Cytotoxin (Paton & Paton, 2010) oder das Cytolethal Distending Toxin V (Janka et al., 2003; Bielaszewska et al., 2005) und Adhärenzfaktoren wie die langen polaren Fimbrien (Jordan et al., 2004; Samadder et al., 2009), Flagellen (Jandhyala et al., 2010; Haiko & Westerlund-Wikström, 2013), das IrgA homologe Adhäsin (Tarr et al., 2000), das STEC Autoagglutinationadhäsin (Paton et al., 2001) oder der hämorrhagische Coli Pilus (Xicohtencatl-Cortes et al., 2007). Außerdem spielen äußere Membranvesikel, durch die Virulenzfaktoren aus der Bakterienzelle geschleust werden können (Kolling & Matthews, 1999; Aldick et al., 2009; Bielaszewska et al., 2017), und eine chemische Kommunikation, das Quorum Sensing (Sperandio et al., 2001; Falcão et al., 2004), eine Rolle bei der Ausprägung der Virulenz.

In den nachfolgenden Kapiteln werden die Hauptvirulenzfaktoren Stx und das T3SS der EHEC genauer beschrieben.

1.2.1 Struktur und Wirkungsweise von Shiga Toxin

Stx gehören zu den bedeutendsten biologischen Toxinen unsere Zeit. Noch heute kommen die veralteten Begriffe „Verotoxin“ oder „Shiga-ähnliches Toxin“ vor. Die toxische Eigenschaft von Stx wird mit Nierenzellen der Grünen Meerkatze, den Verozellen, gemessen, woher sich der Begriff „Verotoxin“ ableitet. „Shiga-ähnliches Toxin“ entstand, weil Shiga Toxin zunächst in *Shigella dysenteriae* entdeckt wurde. Es kommen zwei Stx-Familien vor, die circa 50 % Sequenzübereinstimmung aufweisen, aber immunologisch unterschiedlich sind: Stx1 und Stx2 (Strockbine et al., 1986). Stx1 ist identisch dem *Shigella dysenteriae* Typ I Toxin (O’Brien et al., 1982; Newland et al., 1985). Von Stx1 sind zwei Varianten zum prototypischen Stx1a bekannt, Stx1c und Stx1d, wohingegen Stx2 in den Varianten Stx2a (Prototyp), Stx2b, Stx2c, Stx2d, Stx2e, Stx2f und Stx2g vorkommt (Melton-Celsa, 2014). Stx sind Ho-

lotoxine des AB_5-Typs, die aus einer 32 kDa großen enzymatisch aktiven A-Untereinheit und fünf jeweils 7,7 kDa großen rezeptorbindenden B-Untereinheiten bestehen (Melton-Celsa, 2014). Das *stx* Operon wird von induzierbaren lambdoiden Prophagen oder defekten Prophagen exprimiert (Kapitel 1.3; Scotland et al., 1983; Smith et al., 1983; O'Brien et al., 1984; Unkmeier & Schmidt, 2000). Nach Induktion des entsprechenden *stx*-Prophagens und Lyse der Bakterienzellen wird Stx2 frei (Schmidt, 2001). stx_1 wird durch den Phagenzyklus und einen Eisen-abhängigen Promotor reguliert (Calderwood & Mekalanos, 1987).

Im Darm durchdringt Stx die Epithelbarriere und gelangt in die Blutbahn (Bielaszewska & Karch, 2005; Schüller, 2011). Gebunden an Granulozyten wird es zu den Zielzellen transportiert (te Loo et al., 2000). Das B-Pentamer der meisten Stx-Varianten bindet an den Globotriaosylceramid-Glykosphingolipid Rezeptor an der Oberfläche von Endothelzellen (Jacewicz et al., 1986; Lindberg et al. 1987; O'Brien & Holmes, 1987; Samuel et al., 1990; Johannes & Römer, 2010). Durch Endozytose gelangt das Toxin in die Zelle (O'Brien & Holmes, 1987, Johannes & Römer, 2010, Sandvig et al., 2010). Dort wird es durch einen retrograden Weg vom Endosom, über den Golgi-Apparat zum endoplasmatischen Retikulum transportiert (Johannes & Römer, 2010; Sandvig et al., 2010). Nachdem die A-Untereinheit proteolytisch zum A_1-Fragment reduziert wurde, ist sie für die Inhibierung der Proteinbiosynthese in der Wirtszelle verantwortlich (O'Brien & Holmes, 1987). Das A_1-Fragment agiert als 28S rRNA *N*-Glykosidase und entfernt Adenin vom 3'-Ende der 28S rRNA der 60S Ribosomenuntereinheit (Endo et al., 1988; Saxena et al., 1989). In Folge kommt es zur Apoptose oder Nekrose der humanen Zellen (Bauwens et al., 2011; Amaral et al., 2013). Betroffen sind die mikrovaskulären Endothelzellen der Nieren, sowie Epithel-, Blut- und Nervenzellen (Bielaszewska & Karch, 2005; Tesh, 2010; Betz et al., 2016; Legros et al., 2017).

1.2.2 Der Locus of Enterocyte Effacement

EHEC, enteropathogene *E. coli* (EPEC) und das Mäusepathogen *Citrobacter rodentium* gehören zur Gruppe der A/E-Pathogenen. Die Merkmale von A/E-Läsionen sind die Intimin-vermittelte Anheftungen an die Epithelzellen (Attaching) und die lokale Zerstörung des Mikrovilli Bürstensaums (Effacing) (Jores et al., 2004; Stevens

& Frankel, 2014). Durch Umstrukturierung des Zytoskeletts und Anhäufung von filamentösem Aktin an der Kontaktstelle kommt es zu sockelartigen Ausstülpungen der Zellen, sogenannten Podesten (Jores et al., 2004; Stevens & Frankel, 2014). Für den Phänotyp verantwortlich ist die Expression der Gene der Pathogenitätsinsel Locus of Enterocyte Effacement (LEE), die in fünf polyzistronische Operons unterteilt ist (McDaniel et al., 1995; Perna et al., 1998; Mellies et al., 1999). Der LEE kodiert für ein T3SS (Jarvis et al., 1995), das Adhäsin Intimin (Jerse et al., 1990) und den translozierten Intimin Rezeptor Tir (Kenny et al., 1997), sowie weitere Effektorproteine wie EspG, EspH, EspZ oder Map (Tu et al., 2003; Creuzburg et al., 2011). Darüber hinaus kodiert das LEE1 Operon den Transkriptionsregulator Ler (Mellies et al., 1999; Elliot et al., 2000). Ler aktiviert die Transkription einiger LEE Operons, die unter anderem das T3SS kodieren. Das T3SS ist ein Proteinkomplex mit Nadel-ähnlicher Struktur, die die Zytoplasmamembran, das Periplasma und die äußere Membran der Bakterienzelle durchdringt (Stevens & Frankel, 2014; Gaytán et al., 2016). Nach Kontakt mit der Zielzelle wird in deren Membran eine Pore gebildet, durch die Effektorproteine in die Wirtzelle injiziert werden können. Tir wird an die eukaryotische Zellmembran transloziert und dient dem Adhäsin Intimin als Rezeptor. Weitere Effektorproteine sind an der Manipulation des Signalwegs in der eukaryotischen Wirtszelle beteiligt.

1.3 Shiga Toxin-kodierende Bakteriophagen

In den Genomen von EHEC findet man viele Prophagensequenzen. Ein Prophage ist in das bakterielle Genom integrierte Bakteriophagen (Phagen)-DNA. Die bekanntesten Phagen der EHEC sind ihre Stx-kodierenden Phagen (kurz: Stx-Phagen). Dabei kann ein Bakterienstamm mehrere *stx*-Prophagen in seinem Genom beherbergen. Die *stx*- und andere nicht *stx*-Prophagen von STEC sind für die vorliegende Arbeit von Bedeutung, da die untersuchten Proteine (NanS-p) prophagenkodiert sind (Abbildung 1; Kapitel 1.8.2, 2, 3).

Stx-Phagen und andere EHEC-Phagen sind mit dem Phagen Lambda verwandt und gehören aufgrund ihrer Morphologie zu den Familien *Siphoviridae*, *Myoviridae* oder *Podoviridae* (Schmidt, 2001; Mondal et al., 2016). Phagen verfügen über zwei Rep-

likationsstrategien: den lytischen und den lysogenen Reproduktionszyklus. Während des lytischen Zyklus werden viele infektiöse Phagen gebildet und durch die bakterielle Zelllyse in die Umgebung freigesetzt. Beim lysogenen Zyklus wird die Phagen-DNA in das bakterielle Genom integriert und mit diesem repliziert. Durch Antibiotika, Ultraviolettstrahlung oder weitere Faktoren, wird eine SOS-Antwort ausgelöst und der lytische Zyklus aktiviert. Molekular unterschiedlichste funktionelle und kryptische Stx1- und Stx2-Phagen in verschiedenen STEC-Stämmen wurden beschrieben, deren genetische Ausstattung sequenziert und/oder analysiert. Die prominentesten unter ihnen sind die Stx1-Phagen H-19B aus *E. coli* O26:H11 Stamm H19 (Huang et al., 1987; Neely & Friedman, 1998a & 1998b) und VT1-Sakai aus O157:H7 Stamm Sakai (Yokoyama et al., 2000), sowie die Stx2-Phagen 933W aus O157:H7 Stamm EDL933 (Plunkett et al., 1999) und VT2-Sakai ebenfalls aus *E. coli* Sakai (Makino et al., 1999). Die unterschiedlichen Stx-Phagen sind ein Resultat aus multiplen genetischen Rekombinationen ihrer DNA (Neely & Friedman, 1998a).

1.3.1 Genetische Organisation der Shiga Toxin-Bakteriophagen

Die Stx-Phagen haben ein doppelsträngiges DNA-Genom. Sie integrieren ihre DNA immer in bestimmten Stellen im bakteriellen Chromosom, z.B. *wrbA* (Plunkett et al., 1999), *sbcB* (Ohnishi et al., 2002), *yehV* (Perna et al., 2001) oder *yecE* (Recktenwald & Schmidt, 2002). Die Phagen-DNA kann in eine früh und eine spät regulierte Genregion unterteilt werden. Zur früh regulierten Genregion gehört (i) der *int-xis* Abschnitt, der für die Integration und Exzision der DNA verantwortlich ist, (ii) die Rekombinationsregion mit *exo*, *gam* und *bet,* die für Proteine kodieren, die am Lambda „Red Recombinase“ System beteiligt sind und (iii) die Replikationsregion mit *cII*, *cIII,* die Transkriptionsregulatoren kodieren, dem Antiterminator N Gen, den Regulatorgenen *cro* und *cI*, sowie den Replikationsprotein O und P Genen (Abbildung 1) (Neely & Friedman, 1998a; Tyler et al., 2004; Smith et al., 2012). Die Stx-Gene befindet sich in der spät regulierten Genregion zwischen dem Antiterminator Q Gen und den Genen der Lysekassette (Abbildung 1) (Karch et al., 1999; Neely & Friedman, 1998b; Plunkett et al., 1999; Schmidt, 2001). Zwischen *Q* und *stx* können drei Gene liegen, die für seltene tRNAs kodieren (Plunkett et al., 1999; Schmidt et

al., 1997). Die Lysekassette besteht aus Genen, die für Holin oder Pinholin (*S*), Endolysin (*R*) oder Proteine des Spaninkomplexes (*Rz* und *Rz1*) kodieren (Berry et al., 2008; Savva et al., 2008; Pang et al., 2009; Cahill et al., 2017). Im Lambda-Phagen findet man in der spät regulierten Genregion auch Gene für Verpackung und Morphologie, wie Kapsid und Schwanz (Smith et al., 2012; Mondal et al., 2016). Die meisten Genprodukte der spät regulierten Genregion der Stx-Phagen sind jedoch unbekannt (Mondal et al., 2016). Es gibt einige Ausnahmen, wie der stromabwärts zu *stx* lokalisierte Lokus Z1466 (in dieser Arbeit als *nanS*-p1 deklariert (Kapitel 2)) von EDL933, das eine *O*-Acetyl Esterase kodiert (Nübling et al., 2014) oder neun spät regulierte Gene von Stamm Sakai, deren Genprodukte an der Morphologie des Phagens beteiligt sind (Abbildung 1) (Mondal et al., 2016). *lom* kodiert einen Adhärenzfaktor, der an der Adhärenz an Wangenepithelzellen beteiligt ist (Vica Pacheco et al., 1997) und *bor* kodiert ein Lipoprotein, das für Resistenz in Serum verantwortlich ist (Barondess & Beckwith, 1995). Man geht davon aus, dass viele spät regulierten Gene Nutzen für die Bakterien haben und nicht für den Phagen selbst (Hendrix et al., 2000; Smith et al., 2012).

1.3.2 SOS-Antwort der Wirtszelle und *stx*-Expression

Bakterien reagieren auf extreme DNA-Schäden mit einer SOS-Antwort, um zu überleben (Baharoglu & Mazel, 2014). Regulatoren der SOS-Antwort sind die Proteine LexA und RecA. LexA reprimiert normalerweise die Expression der SOS-Gene und wird durch RecA nach DNA-Beschädigung inaktiviert (Little, 1983). Die SOS-Gene werden exprimiert und DNA-Reparaturen veranlasst (Baharoglu & Mazel, 2014). Das SOS-Signal beeinflusst jedoch auch die temperenten Stx-Phagen, die in den lytischen Zyklus übergehen (Matsushiro et al., 1999; Herold et al., 2005; Nassar et al., 2013). Die regulatorischen Proteine CI und Cro sind für den Erhalt des lysogenen Zustands des Phagens bzw. die Initiation des lytischen Zyklus verantwortlich (Neely & Friedman, 1998a). Der Repressor CI bindet spezifisch an Operatoren rechts und links von *cI* (Abbildung 1) und unterdrückt die Expression der früh regulierten Gene (Ptashne, 1992). *Cro* befindet sich stromabwärts von *cI* (Abbildung 1) und Cro bindet ebenfalls an Operatoren rechts und links von *cI* und ist für die Einleitung des lytischen Zyklus verantwortlich (Ptashne, 1992). RecA inaktiviert den

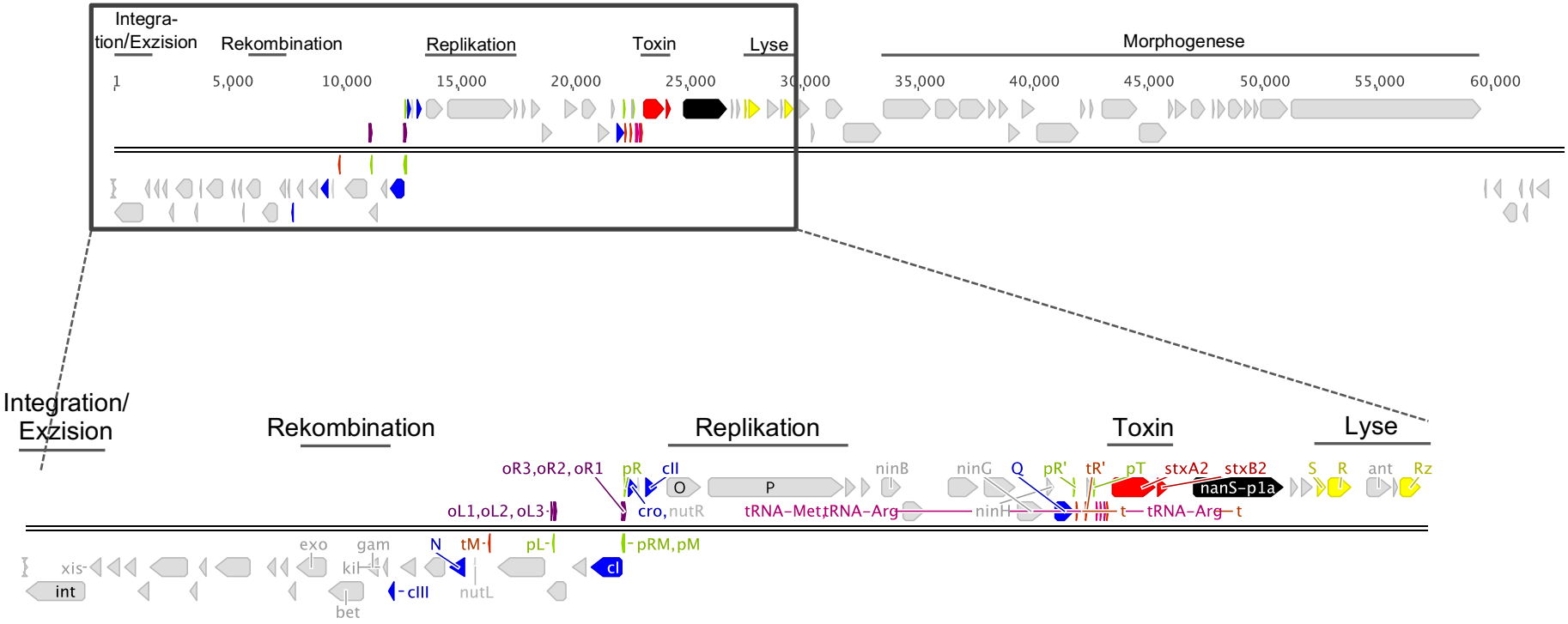

Abbildung 1: *stx*$_{2a}$-Prophage 933W (63132 bp) von *E. coli* O157:H7 Stamm EDL933 (Chromosomensequenz CP015855 (Kapitel 2) hinterlegt bei der Datenbank des National Center for Biotechnology Information (NCBI) (Anonym c, 2018)). Oben befindet sich die gesamte DNA des Phagens; Unten ist die Sequenz zwischen *int* und *Rz* detaillierter dargestellt Blau: Regulationsgene; Violett: Operatorsequenzen; Orange: Terminatorsequenzen; Grün: Promotersequenzen; Pink: tRNA Gene; Rot: *stxA*$_{2a}$, *stxB*$_{2a}$; Schwarz: *nanS*-p1a; Gelb: Lysegene; Grau: Andere oder unbekannte Gene und *nut* Sequenzen. Dargestellt mit Geneious Software Version 9.1.8 (Biomatters Ltd., Auckland, Neuseeland). Identifikation der Gene durch Blastn Analyse (Anonym b, 2018).

Repressor CI und initiiert damit den lytischen Zyklus (Ptashne, 1992; Roca & Cox, 1997). Die Transkription der frühen Gene beginnt an den Promotoren p_R und p_L. Transkripte des Antiterminators N Gens können nun von p_L aus gebildet werden. Außerdem werden Transkripte des Cro Gens von p_R aus gebildet. Der Antiterminator N agieren an RNA *nut* Stellen (Abbildung 1) durch Modifikation der RNA Polymerase, so dass diese die stromabwärts liegenden Terminatoren überliest bis hin zu Q (Das, 1992; Friedman & Court, 1995). Der Antiterminator Q agiert am späten Promoter p'_R und ist verantwortlich für das Überlesen des Terminators t'_R in der spät regulierten Genregion (Neely & Friedman, 1998b). Q bindet an die Erkennungsstelle *qut*, das einen Teil von p_R' und einen stromabwärts liegenden DNA Abschnitt abdeckt (Yarnell & Roberts, 1992). Nach Induktion wird stx_2 von p_R' ausgehend exprimiert (Neely & Friedman, 1998b). stx_1 hat noch einen weiteren eigenen Promotor (Calderwood & Mekalanos, 1987). Durch die Phagenlyse kommt es zum Zelltod und Freisetzung von infektiösen Phagen und Stx.

1.4 Bakterielle Interaktion mit Mukus des Gastrointestinaltrakts

Mukus ist eine beständige viskose Substanz, die die Atemwege, Augen, den Genitaltrakt und die inneren Organe, die im Kontakt mit der Umwelt stehen, bedeckt. Mukus hat die Funktion, das Epithel vor chemischem, mikrobiellem, enzymatischem und mechanischem Schaden zu schützen. Dies wird durch seine hohe Viskosität, die kontinuierliche Abstoßung und Regeneration, der Produktion von Lysozym-ähnlichen Enzymen oder der Möglichkeit des Partikeleinschlusses gewährleistet (Bakshani et al., 2018). Der Mukus des Gastrointestinaltrakts trennt sich in zwei Schichten: eine lose freie (äußere) und eine membrangebundene (innere) Schicht, die mit der Mukosa verbunden ist und ein Bestandteil der Glykokalyx der Zelloberfläche darstellt (Atuma et al., 2001; Johansson et al., 2008; Johansson et al., 2011; Jonckheere et al., 2013). Mukus wird von Gobletzellen gebildet (Johansson & Hansson, 2013; Pelaseyed et al. 2014). Die Dicke der Mukusschicht im Gastrointestinaltrakt variiert und ist im Dickdarm am größten (Atuma et al., 2001).

Muzine sind die Glykoproteine des Mukus. Die Glykane bilden lineare und verzweigte Strukturen aus und bestehen aus unterschiedlich vielen Monomeren, die post-translational angefügt werden. Diese sind Ribose, Arabinose, *N*-Acetylgalakto-

samin (GalNAc), *N*-Acetylglukosamin (GlcNAc), Fukose, Mannose, Galaktose, Glukonat, Glukuronat, Galakturonat, Galaktosamin, Glukosamin, 5-*N*-Acetylneuraminsäure (Neu5Ac) und Derivate dieser Zucker, die die Substituenten Sulfat, Acetat, Acyl, Methyl oder Phosphat enthalten (Pacheco & Sperandio, 2015). Serin oder Threonin bilden mit ihren Hydroxylgruppen glykosidische Bindungen mit dem ersten Monosaccharid der Glykankette (Corfield, 2018). Die Aminosäuren gehören der Prolin-Threonin-Serin Domäne an, die als Tandem Repeat gebildet werden und sich in ihrer Größe und ihrem Muster in den verschiedenen Muzinen unterscheiden (Corfield, 2018). Das terminale Monosaccharid ist Fukose, Galaktose, GalNAc oder Neu5Ac, und gehört zu den Histoblutgruppen Antigenen (Larsson et al., 2009; Moran et al., 2011; Juge, 2012). Vom Dünndarm zum Dickdarm nimmt die Konzentration an endständigen Sialinsäuren zu, während die Konzentration an Fukose abnimmt (Robbe et al., 2003). Seltener kommen auch *N*-verknüpfte Glykane vor, die mit Asparagin verbunden sind (Corfield, 2015; Corfield, 2018). Muzine können in die Kategorien (i) sekretiert Gel-bildend, (ii) sekretiert nicht-Gel-bildend und (iii) Zellmembran-gebunden Muzine eingeteilt werden (Moran et al., 2011; Corfield, 2018). Das sekretierte Gel-bildende MUC2 ist im Dickdarm des Menschen das am häufigsten vorkommende Muzin. Das Epithel ist von den Membran-gebundenen Muzinen MUC1, MUC3, MUC4, MUC12, MUC13, MUC15, MUC17, MUC20 und MUC21 bedeckt (Moran et al., 2011).

Muzine werden durch Glykosidasen von anaeroben Bakterien des Darms konstant zersetzt (Corfield et al., 1992; Conway & Cohen, 2015; Tailford et al., 2015). Einige, wie *Bifidobacterium bifidum*, *Bacteroides thetaiotaomicron, Bacteroides fragilis*, *Akkermansia muciniphila* oder *Ruminococcus gnavus*, sind spezialisiert, Monosaccharide wie Sialinsäure und Fukose aus Muzin freizusetzen und diese als Energiequelle zu nutzen (Macfarlane & Gibson, 1991; Crociani et al., 1994; Xu et al., 2003; Derrien et al., 2004; Crost et al., 2013; Corfield, 2018). Auch nicht-Muzin-spaltende Mitglieder der Darmmikrobiota können die freigesetzten Monosaccharide nutzen (Sonnenburg et al., 2010; Turroni et al., 2015; Arike & Hannson, 2016). *E. coli* können keine komplexen Polysaccharide abbauen und sind auf Mono- oder Disaccharide angewiesen (Hoskins et al., 1985; Conway & Cohen, 2015; Tailford et al., 2015). *E. coli*

und andere Bakterien besiedeln das äußere Muzin, während das innere nahezu steril ist (Johansson et al., 2008 & 2011; Hansson & Johansson, 2010).

1.4.1 Interaktion enteropathogener Bakterien mit Mukus

Intestinale Pathogene wie *Salmonella enterica* Serotyp Typhimurium (*S.* Typhimurium), *Campylobacter jejuni*, *Clostridium difficile*, *Vibrio cholerae* und humanpathogene STEC besitzen Strategien, um Mukus zur Infektion zu nutzen (Sicard et al., 2017; Bakshani et al., 2018). Sie können um Mukus-stämmige Kohlenhydrate konkurrieren, Mukusbestandteile als Rezeptoren verwenden, die Expression von Mukus regulieren oder den inneren Mukus gezielt penetrieren, um zum Epithel vorzudringen (Conway & Cohen, 2015; Sicard et al., 2017; Bakshani et al., 2018). Mukus selbst beeinflusst wiederrum die Expression verschiedener Virulenzgene pathogener Bakterien (Vogt et al., 2015).

Kohlenhydrate aus Mukus im Darm sind eine sich kontinuierlich erneuernde Energiequelle. Sie können von Pathogenen als Wettbewerbssubstrate genutzt werden, um erfolgreich kolonisieren zu können. So konnte im Rattenmodell eine Beteiligung von Mannose aus Muzin an der Adhärenz von *S.* Typhimurium nachgewiesen werden (Vimal et al., 2000). *C. jejuni* sichert sich eine Nische, da sie hauptsächlich Aminosäuren als Kohlenhydratquelle verwenden (Guccione et al., 2008; Hofreuter et al., 2008). Tatsächlich fehlen ihnen viele Gene des Zuckermetabolismus. Lediglich Fukose kann von *C. jejuni* als Monosaccharid katabolisiert werden (Muraoka & Zhang, 2011; Stahl et al., 2011). In *V. cholerae* unterstützte Neu5Ac das frühe Infektionsstadium der Bakterien im Mäusemodell indem es die Konkurrenzfähigkeit von *V. cholerae* beeinflusste (Almagro-Moreno & Boyd, 2009; McDonald et al., 2016; Reddi et al., 2018). *V. cholerae* exprimiert die Sialidase NanH und kann sich Sialinsäuren aus Muzin bereitstellen. Eine besondere Herausforderung beim Wettbewerb um Mukus-stämmige Kohlenhydrate hat STEC, da die nahe verwandten kommensalen *E. coli* im Darm die gleichen Substrate verstoffwechseln (Kapitel 1.6.1). In der vorliegenden Arbeit sollten die Beteiligung von Sialinsäureesterasen, die von einigen enteropathogenen *E. coli* exprimiert werden, im Wettbewerb um Sialinsäuren analysiert werden (Kapitel 3). Speziell Sialinsäurederivate von Neu5Ac standen dabei im Fokus.

Doch nicht nur als Kohlenhydratlieferant wird Mukus von Pathogenen verwendet, sondern auch zur Adhäsion. *S.* Typhimurium zeigte im Zellkulturmodell eine höhere Adhärenz an die mukusbildende Zellline HT29-MTX als an vergleichbare Zellen, die keinen Mukus exprimieren (Gagnon et al., 2013). *C. jejuni*, *C. difficile* und EHEC adhärieren mit ihren Flagellen an Mukus (Tasteyre et al., 2001; Erdem et al., 2007; Young et al., 2007). Das heißt, dass Pathogene Muzinbausteine als Rezeptoren verwenden. Muzin-stämmiges GlcNAc ist beispielsweise der Rezeptor für das Oberflächenprotein GbpA von *V. cholerae* (Bhowmick et al., 2008; Reddi et al., 2018). Ein Exkurs in die Virologie zeigt, dass humanpathogene Viren häufig Sialolektine zur Adhärenz an Mukus nutzen, indem sie Sialinsäuren als Rezeptoren verwenden (Wasik et al., 2017).

Auch die Bildung von Mukus, Muzin oder bestimmter Mukussaccharide kann durch Pathogene reguliert werden. *C. difficile* reduziert die Konzentrationen von GalNAc und MUC2 und erhöht die von GlcNAc und terminaler Galaktose (Engevik et al., 2015). GbpA und das Choleratoxin von *V. cholerae* sind in der Lage, die Sekretion von Muzin zu erhöhen (Lencer et al., 1990; Epple et al., 1997; Bhowmick et al., 2008). *E. coli* O157:H7 erhöhte die Expression von MUC2 in HT-29 Zellkulturinfektionsversuchen (Xue et al., 2014). Auch mit dem EHEC verwandten A/E-Mäusepathogen *Citrobacter rodentium* wurde eine erhöhte Sekretion von MUC2 in Mäusen beobachtet (Bergstrom et al., 2015). Verantwortlich ist der Tumornekrosefaktor α, der von den Wirtszellen als Reaktion gebildet wird (Xue et al., 2014).

Enteropathogene können in den inneren Mukus eindringen, um zum Epithel zu gelangen. Durch seine typisch helikale Struktur wird *C. jejuni* die Passage durch Mukus gewährt (Stahl et al., 2016). Jedoch spielt vor allem die Beweglichkeit und die Anwesenheit von Mukus-degradierenden Enzymen eine Rolle bei der Invasion. *V. cholerae* nutzt die Muzinasen Hap und TagA, um den Mukus des Dünndarms zu penetrieren, so dass die Bakterien das Epithel erreichen (Booth et al., 1983; Silva et al., 2003; Szabady et al., 2011; Almagro-Moreno et al., 2015). *S*. Typhimurium bildet die Sialidase NanH (EC 3.2.1.18) und die Amylase MalS zur Degradierung von Glykanen (Hoyer et al., 1992; Arabyan et al., 2016). *E. coli* O157:H7 kann durch die Muzinase StcE (EC 3.4.24.-) MUC2 gezielt penetrieren (Grys et al., 2005; Grys et al., 2006; In et al., 2016; Hews et al., 2017). Auch Serinproteaseautotransporter

von *Enterobacteriaceae*, wie die von *E. coli* O104:H4 Stamm LB226692 kodierten (Kapitel 1.1), können muzinolytische Eigenschaften haben.
Zuletzt reguliert Mukus oder dessen Bestandteile die Virulenz der Pathogene. Dadurch kann die Transkription der entsprechenden Gene zum optimalen Zeitpunkt und am richtigen Ort stattfinden. Tu et al. (2008) zeigten, dass MUC2 die Expression von Genen hochreguliert, die an Mobilität, Adhäsion, Invasion und Toxinbildung von *C. jejuni* beteiligt sind. In *V. cholerae* erhöht GlcNAc die Mobilität der Bakterien (Bhowmick et al., 2008; Reddi et al., 2018). In EHEC wird die LEE Expression und andere Virulenzfaktoren durch Zucker aus Muzin beeinflusst (Kapitel 1.6.2).

1.5 EHEC-Infektion des humanen Dickdarms

Der Dünn- und vor allem der Dickdarm sind komplexe Ökosystem, die durch eine charakteristische mikrobielle Gemeinschaft besiedelt sind. EHEC infizieren das terminale Ileum und den Dickdarm. Der Dickdarm ist das Habitat der meisten Bakterien im Menschen. Starke Entzündungen nach Infektionen mit *E. coli* O157 in Patienten mit hämorrhagischer Kolitis wurden im ersten Abschnitt des Dickdarms, dem Caecum, und im aufsteigenden Dickdarm gefunden (Shigeno et al., 2002). Im weiteren Verlauf des Dickdarms, über den transversalen Dickdarm, den absteigenden Dickdarm und das Sigmoid, nahmen die Schädigungen des Epithels ab (Shigeno et al., 2002). Auch in infizierten Mäusen oder Kaninchen war vorrangig der Dickdarm betroffen (Li et al., 1993; Karpman et al., 1997; Isogai et al., 1998; Aiba et al., 2002). Vermutlich, weil die initiale Infektionsphase symptomlos verläuft, wurden im Menschen bisher keine A/E-Läsionen nachgewiesen. Symptome treten erst auf, wenn das Epithel bereits zersetzt ist. *In vitro* konnte jedoch für Stämme der Serotypen O157:H7, O103:H^- und O103:H2 Adhärenz und Ausbildung von A/E-Läsionen an Peyer-Plaques an intestinalen Organkulturen des Dünn- oder Zwölffingerdarms beobachtet werden (Phillips et al., 2000; Fitzhenry et al., 2003). Auch Stx-negative Stämme infizierten *ex vivo* zunächst das terminale Ileum und erst später den Dickdarm (Chong et al., 2007). Jedoch wurden nach künstlicher Infektion auch A/E-Läsionen an humanen kolonoiden Monoschichten (In et al., 2016) und an Biopsieproben des Dickdarms beobachtet (Lewis et al., 2015).

Mukus stellt eine Hürde für EHEC dar, die das darunterliegende Epithel kolonisieren. Mit Hilfe von Muzinasen, wie StcE, können sie in den inneren Mukus eindringen (Kapitel 1.4.1). StcE greift zusätzlich noch in die Immunabwehr des Wirts ein. Es wurde gezeigt, dass StcE den Komplement Faktor CD55 auf Zelloberflächen der humanen Karzinomzelllinien HeLa und Caco-2 reduziert (Furniss et al., 2018) und den C1 Esterase Inhibitor verstärkt (Lathem et al., 2002), das eine Inhibierung des Komplementsystems zur Folge hat. Durch die Reduktion von CD55 auf Caco-2 Zellen kommt es zu einer reduzierten Freisetzung von Granulozyten an der apikalen Oberfläche der Zellen (Furniss et al., 2018). Durch beide Mechanismen wird die wirtsspezifische Bekämpfung des Krankheitserregers gehemmt.
Mit der Serinprotease EspP wird die Zerstörung des Mikrovilli Bürstensaums durch Reduktion von Protocadherin 24, einem Zell-Zell Adhäsionsprotein, initiiert (In et al., 2016). Außerdem stimuliert EspP den aktiven Ionentransport (Tse et al., 2018). Im Mäusemodell änderte *E. coli* O157:H7 die Proteinexpression der Tight Junctions, speziell die Expression von Occludin, Claudin-2 und Claudin-3, und erhöhte die intestinale Permeabilität (Roxas et al., 2010). Das Epithel reagiert auf das bakterielle Flagellin und Stx durch eine vermehrte Ausschüttung von Cytokinen und Chemokinen (Isogai et al., 1998; Rogers et al., 2003; Lee et al., 2013; Xue et al., 2014; Lewis et al., 2016; Karve et al., 2017). Ein wichtiger Schritt in der Pathogenese ist die Ausschüttung von Interleukin 8 von Darmepithelzellen (Rogers et al., 2003). Interleukin 8 ist mit der Rekrutierung von Neutrophilen assoziiert (Karve et al., 2017).

1.6 Kontrolle metabolischer und virulenter Genexpression in EHEC durch Kohlenhydrate im Darm

Die Regulierung der Genexpression im Darm ist fein abgestimmt. Während des Infektionsprozesses wird die Nährstoffverfügbarkeit und der physiologische Status des Wirts für die Expression von Metabolismus- und Virulenzgenen von gastrointestinalen Pathogenen herangezogen (Pacheco & Sperandio, 2015). Das macht Sinn, da Energie entweder für den Metabolismus oder für die Ausbildung der Virulenzfaktoren effizient genutzt werden kann. Somit werden der optimale Zeitpunkt und Ort für die Kolonisation festgelegt. An der Virulenzregulierung im Darm generell beteiligt sind Kohlenhydrate, pH-Werte, Gallensalze, Sauerstoffkonzentrationen,

Peristaltik, adrenergene Metaboliten, Quorum Sensing, Ethanolamin und die Darmmikrobiota und deren Metaboliten (Sperandio et al., 2003; Schüller & Pillips, 2010; Pifer & Sperandio 2014; Lewis et al., 2015; Carlson-Banning & Sperandio, 2016; Jubelin et al., 2018). Im Nachfolgenden wird auf den Zuckerkatabolismus und die Virulenzregulierung von EHEC durch Kohlenhydrate im Darm eingegangen.

1.6.1 Verwertung von Kohlenhydraten im Darm

Nach einer Hypothese von Freter ist die Fähigkeit der Bakterien, spezielle Nährstoffquellen effizient nutzen zu können und eine passende Nische zu finden, essentiell für eine erfolgreiche Kolonisation (Freter et al., 1983). Durch metabolische Flexibilität, gemischte Substratverwertung und -umsatz an verschiedenen Orten oder zu unterschiedlichen Zeiten, können metabolisch ähnliche Bakterien jedoch co-existieren (Pereira & Berry, 2017).

EHEC und andere *E. coli* nutzen Mono- und Disaccharide zum Kolonisieren und Persistieren (Pacheco & Sperandio, 2015). Die meisten Zucker aus der menschlichen Ernährung werden im Dünndarm bereits resorbiert und gelangen nicht in den Dickdarm. *E. coli* ist daher auf Zucker aus Ablagerungen des Epithels, Ballaststoffen, Mukus oder auf Intermediate der Darmmikrobiota angewiesen (Miranda et al., 2004; Conway & Cohen, 2015; Pacheco & Sperandio, 2015). Bei der Kolonisation von Streptomycin-behandelten Mäusen durch den kommensalen *E. coli* Stamm MG1655 spielten verschiedene Zuckerstoffwechselwege eine Rolle (Chang et al., 2004). In der frühen Infektionsphase sind vor allem GlcNAc und Neu5Ac wichtig, während in der späteren Phase Glukuronat, Mannose, Fukose und Ribose verstoffwechselt werden (Chang et al., 2004). In diesem Modell konnte EDL933 aufgrund der Zucker Arabinose, Fukose, GlcNAc, Galaktose, Glukuronat, Mannose und Ribose, die alle in Mukus vorkommen, kolonisieren (Fabich et al., 2008). In anderen Studien, die sich mit möglichen Substraten des Rinderdarms auseinander setzten, konnte gezeigt werden, dass Xylose, GalNAc, Ribose, Maltose, Fukose, Mannose, GlcNAc, Neu5Ac und Galaktose an der Vermehrung von *E. coli* O157:H7 beteiligt waren bzw. deren metabolische Stoffwechselwege hochregulierten (Snider et al., 2009; Bertin et al., 2013; Segura et al., 2018). Miranda et al. (2004) fanden heraus,

dass EDL933 in Coinfektionsversuchen mit MG1655 neben glykolytischen auch gluconeogenetische Substrate in der späten Infektionsphase nutzen. MG1655 befand sich im Darmlumen der Mäuse während EDL933 nahe am Epithel lokalisiert war. Im Mukus waren beide Stämme vertreten (Miranda et al., 2004). Schlussendlich besitzen pathogene und kommensale *E. coli* jedoch ein sehr ähnliches Substratspektrum. Für diese Arbeit von besonderer Bedeutung ist der Katabolismus von Sialinsäuren in *E. coli*. Es stellte sich die Frage, ob sich der Sialinsäurekatabolismus von kommensalen *E. coli* von dem von pathogenen *E. coli* unterscheidet (Kapitel 2, 3, 4). In Kapitel 1.8 wird nochmal gesondert auf den Sialinsäurekatabolismus von *E. coli* eingegangen.

1.6.2 Virulenzregulation durch Muzin-stämmige Kohlenhydrate

Ein pathogener Bakterienstamm ist für seinen Wirt nur dann pathogen, wenn er am Infektionsort entsprechende Virulenzfaktoren ausgebildet hat. Für die Regulation sind hauptsächlich Bestandteile des Milieus verantwortlich.

Untersuchungen zur Kohlenhydrat-abhängigen Virulenzregulation in EHEC sind vor allem für den LEE vorhanden. Abhängig von den Transkriptionsfaktoren KdpE und Cra inhibiert ein glykolytisches Milieu die Expression des LEEs, wohingegen ein gluconeogenetisches dies aktiviert (Njoroge et al., 2012). Das heißt, abhängig von den anwesenden Substraten wird Energie durch den Vorgang der Glykolyse oder der Gluconeogenese gewonnen und die Transkription des LEEs entsprechend runter- bzw. hochreguliert. Im Lumen sind glykolytische Substrate vorhanden, am Epithel eher gluconeogenetische, sodass der LEE am geeigneten Ort exprimiert wird (Pacheco & Sperandio, 2015). Hohe Konzentrationen an GlcNAc oder Neu5Ac reprimieren den LEE mit Hilfe des Transkriptionsrepressors NagC des GlcNAc Metabolismus, und der Zuckerkatabolismus wird aktiviert (Le Bihan et al., 2015; Le Bihan et al., 2017). Fukose ist ein ineffizientes Wettbewerbssubstrat für EHEC, da auch kommensale *E. coli* Stämme Fukose katabolisieren (Fabich et al., 2008; Fox et al., 2009; Pacheco & Sperandio, 2015). Bei der Virulenzregulation von *E. coli* O157:H7 ist Fukose jedoch ein wichtiger Initiator (Pacheco et al., 2012). FusK, eine Histidin-Sensorkinase, gehört zum FusKR Zwei-Komponenten System, bei dem FusR ein Regulator darstellt. Durch *B. thetaiotaomicron* wird Fukose von Glykanen

im Darm gespalten. Die höhere Fukoseverfügbarkeit hat eine Repression der LEE-Gene von O157:H7 über den FusKR Signalweg zur Folge (Pacheco et al., 2012). In einer anderen Studie wurde jedoch eine erhöhte LEE-Expression nach Co-Inkubation von *B. thetaiotaomicron* mit EHEC in Mukus detektiert (Inversen et al., 2015). Diese Arbeitsgruppe verwendete für ihre Experimente jedoch FusKR-negative O103 Stämme (Inversen et al., 2015). Die Regulierung der Virulenzfaktoren durch Kohlenhydrate oder direkten Interspezieskontakt ist also fein abgestimmt und Stamm-spezifisch. Über die Regulierung anderer Virulenzfaktoren durch Kohlenhydrate ist wenig bekannt. Die Expression von Stx wurde in den Co-Inkubationsversuchen von Inversen et al. (2015) inhibiert. Mit Transkriptomanalysen konnte eine negative Regulierung von Flagellin durch Komponenten des Muzins in EDL933 identifiziert werden (Kim et al., 2012).

1.7 Struktur von Sialinsäuren und Vorkommen im Menschen

Sialinsäuren gehören zu den α-Ketozuckern und Aminozuckern und bestehen aus neun Kohlenstoffatomen (9-C). Ihnen gehören mehr als 50 verschiedene Derivate an. Gunnar Blix und Ernst Klenk benannten sie als „Sialinsäuren“ vom griechischen Begriff „Sialos (Speichel)“, oder „Neuraminsäuren“ von „Neuronen“ (Lundblad, 2015; Varki et al., 2017). Allen Sialinsäuren gemein ist eine negativ geladene Carboxylgruppe am anomeren zweiten Kohlenstoffatom (C-2) und eine exozyklische Seitenkette an C-6, die 3-C lang ist (C-7 bis C-9). Unterschiedliche Substituenten an C-5 führen mit einem *N*-Acetyl-Rest zur Neu5Ac, mit einem *N*-Glykolyl-Rest zur 5-*N*-Glykolylneuraminsäure (Neu5Gc), mit einem Hydroxyl-Rest zur 2-keto-Deoxynonulosonicsäure (Kdn) oder mit einer freien Aminogruppe zur Neuraminsäure (Neu). Weitere Substituenten können sich an den Hydroxylgruppen von C-4, C-7, C-8 und C-9 befinden. Diese sind Acetyl-, Methyl-, Sulfat- oder Lactyl-Gruppen. Neu5Ac kommt im Menschen am häufigsten vor, während Neu5Gc vom Menschen nicht synthetisiert werden kann (Lundblad, 2015; Varki et al., 2017). Für die Abwesenheit von Neu5Gc im Menschen ist eine Exondeletionsmutation im *CMAH* Gen verantwortlich (Varki, 2009; Springer et al., 2014). *CMAH* kodiert für eine Hydroxylase, die Cytidin-5‘-Monophospho-Neu5Ac (CMP-Neu5Ac) zu CMP-Neu5Gc überführt. Zum Teil wird Neu5Gc dennoch in gesundem Gewebe und in mutiertem

Tumorgewebe gefunden (Pearce & Läubli, 2016; Varki et al., 2017). Ein Zusammenhang besteht vermutlich mit dem Verzehr von rotem Fleisch (Samraj et al., 2015).

Die Biosynthese der verschiedenen Sialinsäurederivate geht immer von Kdn oder Neu5Ac aus. Kdn und Neu5Ac entstehen durch Kondensation von Mannosamin-6-Phosphat bzw. *N*-Acetylmannosamin-6-Phosphat mit Phosphoenolpyruvat im Zytoplasma der eukaryotischen Zelle. Im Nukleus werden sie zu aktivem CMP-Kdn bzw. CMP-Neu5Ac konvertiert, im Golgi-Apparat an Lipid- oder Proteinglykane transferiert und anschließend an die Zelloberfläche transportiert (Varki et al., 2017). Modifikationen entstehen enzymatisch im Golgi-Apparat. *O*-Acetyl an C-9 entsteht jedoch vermutlich durch nicht-enzymatische Migration ausgehend von C-7 (Varki et al., 2017). In den Glykanketten sind Sialinsäuren durch ihr α-konfiguriertes C-Atom an C-2 mit C-3 (α2,3) verknüpft, α2,6 verknüpft mit Galaktose oder *N*-Acetylgalaktosamin oder α2,8 verknüpft mit anderen Sialinsäuren und können so die seltenere Polysialinsäure (polySia) bilden (Mühlenhoff et al., 1998; Varki et al., 2017). Sialinsäuren oder Oligosialinsäuren befinden sich endständig glykosidisch gebunden an Glykokonjugaten, wie *N*-Glykane, *O*-Glykane oder Glykosphingolipide an den Zelloberflächen und geben diesen eine individuelle Maske. In der Leber kommen vor allem Glykoproteine vor, während im Gehirn vor allem Sialoglykolipide (Ganglioside) und polySia vorkommen (Varki et al., 2017; Schnaar et al., 2014). Zur Varianz der Glykanstrukturen kommt es durch Kombination unterschiedlich glykosidisch gebundener Sialinsäuren und deren Derivate (Varki et al., 2017). Das Sialom (= Gesamtheit der Sialinsäuren) unterscheidet sich innerhalb der Zelltypen und innerhalb von Domänen bestimmter Zelloberflächen. Sialinsäuren kommen auch in Lymphozyten, Erythrozyten, lysosomalen Membranen und Blutproteinen vor. Tumorzellen können anhand eines veränderten Sialysierungsmusters identifiziert werden (Shi et al., 1996; Shen et al., 2004; Mandal et al., 2012; Pearce & Läubli, 2016). Auch in humaner Milch befinden sich Sialinsäure-reiche Oligosaccharide, die die Entwicklung der Darmmikrobiota und die Gehirnentwicklung unterstützen und das Neugeborene vor Pathogenen schützen (Lundblad, 2015).

Generell haben Sialinsäuren eine funktionelle Rolle bei Glykan-Protein-, Zell-Zell- oder Pathogen-Zell-Erkennung und stabilisieren Moleküle und Membranen, beteiligen sich an der Immunantwort, Erhöhen die Affinität von Rezeptoren, sind an Signalweitergabe, Wachstum und Differenzierung beteiligt und maskieren Erkennungsstellen wie Rezeptoren oder Galaktose und schützen vor Makrophagen und Hepatozyten (Varki & Gagneux, 2012; Lundblad, 2015; Varki et al., 2017). Sialinsäuren schützen vor Proteasen und Glykosidasen und stellen einen Schutz vor Autoimmunerkrankungen dar (Varki et al., 2017). Andererseits sind sie selbst Liganden für mikrobielle und humane Lektine und können von Pathogenen als Rezeptoren verwendet werden (Lundblad, 2015; Varki et al., 2017). Desweiteren sind sie Kohlenhydrat- und Stickstoffquellen für kommensale und pathogene Mikroorganismen wie *E. coli* (Kapitel 1.8). Eine wichtige Frage dieser Arbeit war, in wie weit Sialinsäurederivate vom Glykoprotein Muzin Bestandteil einer Nährstoffnische für EHEC und anderen STEC sind (Kapitel 3 und 4).

1.8 Katabolismus von Sialinsäuren in *Escherichia coli*

Gene zum Abbau von Sialinsäuren sind in verschiedenen pathogenen und kommensalen Bakterien von Säugetieren vorhanden (Almagro-Moreno & Boyd, 2009). Vermutlich haben die Bakterien dadurch einen selektiven Vorteil und können den Wirt kolonisieren (Almagro-Moreno & Boyd, 2009; Vimr, 2013). Zwischen den verschiedenen Spezies sind die Genanordnungen variabel und es existieren unterschiedliche Transportersysteme (Almagro-Moreno & Boyd, 2009). Auch wenn ihnen Sialidasen fehlen, scheinen *E. coli* Spezialisten unter den Sialinsäure-Verwertern zu sein, denn sie exprimieren mehrere relevante Operons des Sialinsäurekatabolismus (Vimr, 2013).

1.8.1 Genetische Organisation der Sialinsäureverwertung in *E. coli*

Für den Katabolismus von Neu5Ac, Neu5Gc und Kdn oder Derivate dieser Sialinsäuren kodiert das *E. coli* Chromosom für mehrere Enzyme, Regulatoren und Transporter. Für die vorliegende Arbeit relevant ist der Katabolismus von Neu5Ac und *O*-acetylierte Derivate von Neu5Ac. Zur Energiegewinnung aus Neu5Ac wird das *nanATEK-yhcH* und das *nagBAC-umpH* Operon exprimiert. Das *nanATEK-*

yhcH Operon wird durch den Transkriptionsrepressor NanR reguliert (Kalivoda et al., 2003). Dieser binden mit zwei oder mehreren Homodimeren an den Operator, der den *nanA* Promotor überlappt (Kalivoda et al., 2003). NanR wird von *nanR* stromaufwärts des *nanATEK-yhcH* kodiert. Neu5Ac inaktiviert NanR durch Konvertierung des Dimers in eine monomere Form (Kalivoda et al., 2013). Die Proteine die vom *nagBAC-umpH* Operon kodiert werden sind normalerweise am Katabolismus von GlcNAc beteiligt. Jedoch ist GlcNAc-6-Phosphat (GlcNAc-6-P) ein Metabolit des Neu5Ac-Katabolismus (Abbildung 2) und so erklärt sich die Beteiligung dieser Proteine an der Energiegewinnung aus Neu5Ac.

In *E. coli* transportieren die äußeren Membranproteine OmpC, OmpF, NanC und die Permease NanT der Cytoplasmamembran Neu5Ac durch die Zellwand (Abbildung 2) (Vimr & Troy, 1985; Martinez et al., 1995; Hopkins et al., 2013; Vimr, 2013). *O*-acetylierte Derivate dieser Sialinsäuren können nicht durch NanT transportiert werden. Sie müssen zunächst deacetyliert werden (siehe unten). Die Akkumulation von Neu5Ac im Zytoplasma ist zelltoxisch und aktiviert die Aldolase NanA (Vimr & Troy, 1985). NanA spaltet Neu5Ac reversibel zu *N*-Acetylmannosamin (ManNAc) und Pyruvat (Abbildung 2) und wirkt dabei detoxifizierend (Vimr & Troy, 1985; Plumbridge & Vimr 1999). Pyruvat wird im Citratzyklus zu Energie umgesetzt und ManNAc wird durch die ManNAc-Kinase NanK weiter zu ManNAc-6-Phosphat (ManNAc-6-P) phosphoryliert (Abbildung 2) (Plumbridge & Vimr 1999; Vimr, 2013). NanE ist eine ManNAc-6-P 2-Epimerase die ManNAc-6-P zu GlcNAc-6-P konvertiert (Abbildung 2) (Plumbridge & Vimr 1999). NanT, NanA, NanK und NanE werden vom *nanATEK-yhcH* Operon kodiert (Vimr, 2013). Die Funktion von YhcH ist unbekannt. NagA katalysieren die Reaktion von GlcNAc-6-P zu Glukosamin-6-Phosphat (Glc-6-P) und Acetat (Abbildung 2) (Peri et al., 1990; Souza et al., 1997). Die Glc-6-P Deaminase NagB katalysiert die Isomerizierung und Deaminierung von Glc-6-P zu Fruktose-6-P (Fru-6-P) und Ammonium (Abbildung 2) (Plumbridge & Vimr 1999). Die Transkription von *nagA* und *nagB* steht unter der Kontrolle von NagC, das durch GlcNAc-6-P inaktiviert wird (Abbildung 2). *umpH*, das letzte Gen des *nagBAC-umpH* Operons kodiert eine Phosphatase. Diese ist jedoch am Abbau von Neu5Ac nicht beteiligt. Der Katabolismus von Neu5Ac ist mit der Umsetzung zu Fru-6-P beendet. Fru-6-P tritt in die Glykolyse der Zelle ein.

Neben dem *nanATEK-yhcH* kommen in *E. coli* häufig noch zwei weitere NanR regulierte Operone vor: *nanCMS* und *yjhBC* (Vimr et al., 2004; Kalivoda et al., 2013). Das *nanCMS* Operon ist wahrscheinlich für den Katabolismus von Neu5Ac-Derivaten verantwortlich (Vimr, 2013). NanC ist ein Membrankanalprotein (Condemine et al., 2005) und NanM ist eine Neu5Ac-Mutarotase, die α-Neu5Ac zum β-Anomer konvertiert (Severi et al., 2008). In Glykoproteinen kommt α-Neu5Ac vor, dass nach Abspaltung durch Sialidasen spontan zu β-Neu5Ac mutarotiert. β-Neu5Ac wird von

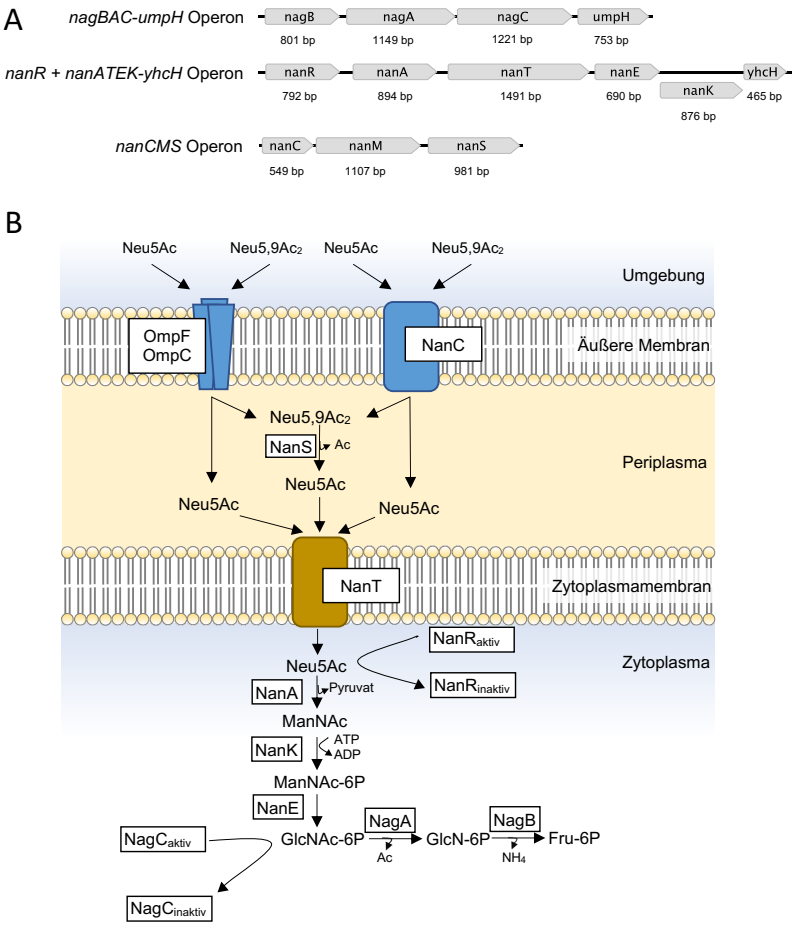

Abbildung 2: Neu5Ac- und Neu5,9Ac$_2$-Katabolismus in *E. coli.* Die Enzyme, Transporter und Regulatoren werden von drei Operons und *nanR* kodiert (NCBI Chromosomensequenzen CP015855 von EDL933 (Kapitel 2); Darstellung mit der Geneious Software) (A). Transport und Katabolismus von Neu5Ac und Neu5,9Ac$_2$ in *E. coli* (B). Ac=Acetat.

E. coli zur Energiegewinnung genutzt. NanM beschleunigt diese Mutarotation (Severi et al., 2008). Durch NanM ist es Sialidase-negativen Bakterien wahrscheinlich möglich, erfolgreich um limitiertes Neu5Ac zu konkurrieren. NanS ist eine *O*-Acetyl Esterase und katalysiert die Hydrolyse von Neu5Ac-Derivaten zu Neu5Ac (Abbildung 2) (Steenbergen et al., 2009; Rangarajan et al., 2011). NanM und NanS kommen vermutlich im Periplasma der Zellen vor (Severi et al., 2008; Steenbergen et al., 2009). Die Funktionen von NanS und NanS-Homologen sind für die vorliegende Arbeit von Bedeutung und deshalb wird in Kapitel 1.8.2 auf die Funktion im Detail eingegangen. Über die Genprodukte des *yjhBC* Operons ist nur wenig bekannt. YjhB ist vermutlich ein Transporterprotein und YjhC ist eine Oxidoreduktase.

1.8.2 Die *O*-Acetyl Esterasen NanS und 933Wp42

NanS ist eine chromosomal-kodierte *O*-Acetyl Esterase, die von vielen *E. coli* Stämmen exprimiert wird (Steenbergen et al., 2009; Rangarajan et al., 2011). Im Chromosom von EDL933 ist *nanS* mit einer Größe von 981 bp präsent (Perna et al., 2001; Latif et al., 2014). Durch die Aktivität von NanS werden Neu5,9Ac_2 und Neu5,8Ac_2 zu Neu5Ac und Acetat umgesetzt (Steenbergen et al., 2009). NanS ist maßgeblich an der Energiegewinnung aus diesen Substraten beteiligt. NanS ist ein Monomer und gehört zur Gruppe II der SGNH Hydrolasen (Rangarajan et al., 2011). Im Prophagen 933W von EDL933, unmittelbar stromabwärts des stx_{2a} befindet sich der Genlokus Z1466 (Perna et al., 2001). Z1466 ist 1938 bp groß und kodiert die 68 kDa große NanS-homologe *O*-Acetyl Esterase 933Wp42 (Nübling et al., 2014). 933Wp42 bewirkt die Freisetzung von Acetat aus Rinderspeicheldrüsenmuzin und Neu5,9Ac_2, und hat ein Temperatur- und pH-Optimum von 50 °C bzw. 7-8 (Nübling et al., 2014). Nach Induktion des Prophagens werden stx_{2a} und Z1466 wahrscheinlich co-transkribiert (Herold et al., 2005). In Proteomanalysen von Polzin et al. (2013) wurde 933Wp42 intrazellulär 40-fach überexprimiert vorgefunden, wenn die Bakterien in einem dickdarmsimulierenden Medium kultiviert wurden. Durch die Lokalisation von Z1466 in direkter Nachbarschaft zu stx_{2a} vor den Lysegenen kann spekuliert werden, dass 933Wp42, wie Stx2a, nach Induktion des Prophagens und Bakterienlyse in den extrazellulären Raum gelangt. Vimr (2013) untersuchte die Anwesenheit des *nanS* Gens in verschiedenen pathogenen und apathogenen *E. coli*

Stämmen. Besonders in Chromosomen von Stämmen, die hämorrhagische Erkrankungen auslösen, sind eine Vielzahl von *nanS*-homologen Genen vorhanden. Sie wurden von Vimr (2013) als kurze, lange, fragmentierte, mittellange, partielle oder sehr lange Kopien von *nanS* beschrieben. Die kurze Kopie entspricht dem eigentlichen *nanS*. Die Gene können in 5'-und 3'-Richtung verlängert sein. Eine weitere Besonderheit ist ihre Lokalisation in Prophagen (Vimr, 2013). Auch Z1466 hat neben der *nanS*-homologen Sequenz weitere DNA Abschnitte und wurde von Vimr (2013) unter „lange *nanS* Kopie" eingeordnet. Laut Vimr (2013) sind im Chromosom von EDL933 eine kurze, 5 partielle und 7 lange *nanS* Kopien vorhanden. Im EHEC O26:H11 Chromosom sind es sogar eine kurze und 11 lange *nanS* Kopien. Der enteroaggregative *E. coli* Stamm 55989 hat keine kurze, aber 3 lange *nanS* Kopien. Die meisten der untersuchten Stämme besitzen eine kurze *nanS* Kopie, und nur wenige haben keine oder zwei kurze *nanS* Kopien. Auffallend ist, dass besonders viele lange *nanS* Kopien in Chromosomen von EHEC vorkommen. Im Rahmen dieser Arbeit wurden die „*nanS* Kopien" als „*nanS*-p" deklariert und Z1466 und 933Wp42 entsprechend als *nanS*-p1 bzw. NanS-p1 (Kapitel 2). „p" steht dabei für „Phage". Welche Funktion die multiplen NanS-p Proteine für das Bakterium haben war zum Beginn der Arbeit nicht bekannt.

1.9 Zielsetzung der Arbeit

In den Chromosomen von EHEC und anderen pathogenen *E. coli* findet man multiple *nanS*-p Gene. Über die Funktion und Bedeutung dieser *nanS*-p Gene und der entsprechenden putativen Proteine war zum Zeitpunkt des Beginns der Arbeit nur wenig bekannt. Im Rahmen dieser Arbeit sollte die Verteilung der *nanS*-p Gene, einige enzymatischen Eigenschaften der entsprechenden NanS-p Proteine, ihre Substratspezifität und die Interaktion mit Mitgliedern der Darmmikrobiota analysiert werden. Die Frage ist, warum besitzen EHEC multiple *nanS*-p Gene und kodieren diese für funktionelle Proteine? Und welche Bedeutung haben NanS-p Proteine bei der Bereitstellung von Substraten? Bewirken sie ein erweitertes Substratspektrum von pathogenen *E. coli* gegenüber kommensalen *E. coli* oder können die Bakterien durch die Anwesenheit von NanS-p eine Nährstoffnische effizienter nutzen? Es sollte ein Modell vorgestellt werden wie EHEC *in vivo* um Nischen zur Kolonisation

im Darm konkurrieren könnten. Zur Beantwortung sollten bioinformatische, molekularbiologische, biochemische und mikrobielle Verfahren zum Einsatz kommen. Die Fragestellung war Teil eines DFG Projektes (Schm1360/6-1). Die erzielten Ergebnisse wurden in drei wissenschaftlichen Publikationen (Kapitel 2-4) veröffentlicht.

In Kapitel 2 werden Ergebnisse zur Quantität der *nanS*-p Gene im Chromosom, zum Auftreten von Domänen in ihren entsprechenden Aminosäuresequenzen und zur Lokalisation der Gene in *E. coli* O157:H7 Stamm EDL933 vorgestellt. *nanS* und alle *nanS*-p Gene im EDL933 Chromosom wurden deletiert. Wachstumskurven der entsprechenden Deletionsmutantenstämme mit Neu5,9Ac_2 als Kohlenhydratquelle werden dargestellt und verdeutlichen die Bedeutung von NanS und NanS-p Proteinen beim Katabolismus von Neu5,9Ac_2. Ergebnisse zu ausgewählten rekombinant hergestellten NanS-p Proteinen, einige enzymatische Charakteristika und Ergebnisse zum extrazellulären Einsatz der NanS-p Proteine in den Kultivierungsmedien der EDL933 Deletionsmutantenstämme und einer *nanS*-Deletionsmutanten des kommensalen *E. coli* Stamm C600 werden gezeigt.

In Kapitel 3 werden die Daten der *in silico* Untersuchungen der Chromosomensequenz des LB226692 Stammes und anderen Chromosomensequenzen, die bei NCBI hinterlegten sind, präsentiert. Sukzessive Deletionen der *nanS*-p Gene im Chromosom des O104:H4 Stammes C227-11φcu und Untersuchungen zu einem potenziellen Gen-Dosis Effekt wurden durchgeführt. Um die NanS-p abhängige Interaktion der STEC mit kommensalen *E. coli* zu untersuchen, wurden Co-Kultivierungen zwischen C227-11φcu oder dessen Deletionsmutantenstämmen und dem kommensalen *E. coli* Stamm AMC 198 mit Neu5,9Ac_2 oder Rinderspeicheldrüsenmuzin durchgeführt.

In Kapitel 4 werden Ergebnisse zu zwei Themen präsentiert: das NanS/NanS-p Substratspektrum (bearbeitet von Stefanie Feuerbaum und Gottfried Pohlentz) und der potenzielle synergistische Effekt von NanS-p Proteinen und einer Sialidase (bearbeitet von Nadja Saile). Bestandteil der vorliegenden Dissertation ist der von Nadja Saile bearbeitete Teil. Es werden Messwerte einer optischen Trübungsmessung nach Kultivierung des EDL933 Stammes und entsprechender Deletionsmutantenstämmen mit einer rekombinant hergestellten Sialidase aus *B. thetaiotaomicron* (BTSA) in Rinderspeicheldrüsenmuzin dargestellt.

Kapitel 2

Escherichia coli O157:H7 strain EDL933 harbors multiple functional prophage-associated genes necessary for the utilization of 5-*N*-acetyl-9-*O*-acetyl neuraminic acid as a growth substrate.

Nadja Saile[1], Anja Voigt[1], Sarah Kessler[1], Timo Stressler[2], Jochen Klumpp[3], Lutz Fischer[2], Herbert Schmidt[1]

[1] Department of Food Microbiology and Hygiene, Institute of Food Science and Biotechnology, University of Hohenheim, Stuttgart, Germany

[2] Department of Biotechnology and Enzyme Science, Institute of Food Science and Biotechnology, University of Hohenheim, Stuttgart, Germany

[3] Department of Health Sciences and Technology, Institute of Food, Nutrition and Health, ETH Zurich, Zurich, Switzerland

Publiziert in Applied and Environmental Microbiology, Band 82, Seiten 5940-5950.

DOI: 10.1128/AEM.01671-16. Mit freundlicher Genehmigung von American Society for Microbiology.

Diese Publikation beinhaltet eine zusätzliche Tabelle mit Primersequenzen und PCR Programmen (Supplemental material), die unter dem Link http://dx.doi.org/10.1128/AEM.01671-16 zu finden ist.

AMERICAN SOCIETY FOR MICROBIOLOGY | Applied and Environmental Microbiology

Escherichia coli O157:H7 Strain EDL933 Harbors Multiple Functional Prophage-Associated Genes Necessary for the Utilization of 5-*N*-Acetyl-9-*O*-Acetyl Neuraminic Acid as a Growth Substrate

Nadja Saile,[a] Anja Voigt,[a] Sarah Kessler,[a] Timo Stressler,[b] Jochen Klumpp,[c] Lutz Fischer,[b] Herbert Schmidt[a]

Department of Food Microbiology and Hygiene[a] and Department of Biotechnology and Enzyme Science,[b] Institute of Food Science and Biotechnology, University of Hohenheim, Stuttgart, Germany; Department of Health Sciences and Technology, Institute of Food, Nutrition and Health, ETH Zurich, Zurich, Switzerland[c]

ABSTRACT

Enterohemorrhagic *Escherichia coli* (EHEC) O157:H7 strain EDL933 harbors multiple prophage-associated open reading frames (ORFs) in its genome which are highly homologous to the chromosomal *nanS* gene. The latter is part of the *nanCMS* operon, which is present in most *E. coli* strains and encodes an esterase which is responsible for the monodeacetylation of 5-*N*-acetyl-9-*O*-acetyl neuraminic acid (Neu5,9Ac$_2$). Whereas one prophage-borne ORF (z1466) has been characterized in previous studies, the functions of the other *nanS*-homologous ORFs are unknown. In the current study, the *nanS*-homologous ORFs of EDL933 were initially studied *in silico*. Due to their homology to the chromosomal *nanS* gene and their location in prophage genomes, we designated them *nanS*-p and numbered the different *nanS*-p alleles consecutively from 1 to 10. The two alleles *nanS*-p2 and *nanS*-p4 were selected for production of recombinant proteins, their enzymatic activities were investigated, and differences in their temperature optima were found. Furthermore, a function of these enzymes in substrate utilization could be demonstrated using an *E. coli* C600Δ*nanS* mutant in a growth medium with Neu5,9Ac$_2$ as the carbon source and supplementation with the different recombinant NanS-p proteins. Moreover, generation of sequential deletions of all *nanS*-p alleles in strain EDL933 and subsequent growth experiments demonstrated a gene dose effect on the utilization of Neu5,9Ac$_2$. Since Neu5,9Ac$_2$ is an important component of human and animal gut mucus and since the nutrient availability in the large intestine is limited, we hypothesize that the presence of multiple Neu5,9Ac$_2$ esterases provides them a nutrient supply under certain conditions in the large intestine, even if particular prophages are lost.

IMPORTANCE

In this study, a group of homologous prophage-borne *nanS*-p alleles and two of the corresponding enzymes of enterohemorrhagic *E. coli* (EHEC) O157:H7 strain EDL933 that may be important to provide alternative genes for substrate utilization were characterized.

Enterohemorrhagic *Escherichia coli* (EHEC) bacteria are foodborne pathogens that cause severe human gastrointestinal illness characterized by bloody diarrhea and hemolytic-uremic syndrome (HUS) (1, 2). EHEC are very heterogeneous in their genome sizes and structures as well as in their virulence gene composition (3, 4). The major pathogenicity factors of strains of classical EHEC serogroups such as O157, O26, O111, O145, and O103 are production of one or more Shiga toxins (Stx) and development of attaching and effacing lesions in the human large intestine (5). The latter effect is caused by the translocation of effector proteins into eukaryotic cells by a type III secretion system encoded by the locus of enterocyte effacement (LEE) (6). Besides LEE-positive EHEC strains, LEE-negative EHEC strains such as O113:H21 strain 98NK2 and the O104:H4 clone have caused severe human disease and outbreaks (7, 8). Stx are generally encoded in the genome of lambdoid phages. EHEC strains can carry one or more Shiga toxin-encoding prophages (Stx prophages) in their genome (9, 10). Moreover, a number of functional and cryptic non-Stx prophages have been identified in the EHEC genome. Stx phages generally carry their *stx* genes in a distinct DNA region downstream of the *Q* antiterminator and upstream of the *S* lysis gene, causing production of Stx during expression of the late phage genes and release of Stx during phage-mediated cell damage (11, 12).

In the same DNA region, a large open reading frame (ORF) has been described in many Stx2a-encoding phages (9). In a study by Herold et al. (13), it was shown that the corresponding ORF of phage 933W and other phages was expressed upon norfloxacin induction. Later on, the corresponding 933Wp42 protein was identified as the most prominently regulated protein, with 40-fold overexpression in simulated colonic environmental medium (SCEM) under aerobic conditions in strain EDL933 (14). Consequently, protein 933Wp42 was recombinantly produced from laboratory *E. coli* strain C600/933W and demonstrated cleavage of 5-*N*-acetyl-9-*O*-acetyl neuramininic acid (Neu5,9Ac$_2$) (15). In a

Received 1 June 2016 **Accepted** 14 July 2016

Accepted manuscript posted online 29 July 2016

Citation Saile N, Voigt A, Kessler S, Stressler T, Klumpp J, Fischer L, Schmidt H. 2016. *Escherichia coli* O157:H7 strain EDL933 harbors multiple functional prophage-associated genes necessary for the utilization of 5-*N*-acetyl-9-*O*-acetyl neuraminic acid as a growth substrate. Appl Environ Microbiol 82:5940–5950. doi:10.1128/AEM.01671-16.

Editor: C. A. Elkins, FDA Center for Food Safety and Applied Nutrition

Address correspondence to Herbert Schmidt, herbert.schmidt@uni-hohenheim.de.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/AEM.01671-16.

TABLE 1 Characteristics of *nanS* and *nanS*-p1-p10 and corresponding proteins in *Escherichia coli* O157:H7 strain EDL933 (accession no. NZ_CP008957.1)[a]

Phylogroup	Gene designation	Gene length (bp)	Chromosomal position in EDL933 (NZ_CP008957)	Designation of corresponding protein	Protein length (aa)
	nanS	981	5440406–5441386	NanS	326
1	*nanS*-p1 (z1466)	*1,938*	*1356672–1358609*	NanS-p1 (*933Wp42*)	*645*
1	*nanS*-p2	1,947	3010352–3012298	NanS-p2	648
3	*nanS*-p3	1,851	1267489–1269339	*	616
3	*nanS*-p4	1,851	2332008–2333858	NanS-p4	616
3	*nanS*-p5	1,854	3571915–3573768	*	617
3	*nanS*-p6	1,854	1859884–1861737	*	617
3	*nanS*-p7	1,851	2789290–2791140	*	616
2	*nanS*-p8	**1,938**	**1642848–1644785**	*	**645**
3	*nanS*-p9	**1,851**	**1907347–1909197**	*	**616**
3	*nanS*-p10	**1,854**	**2152546–2154399**	*	**617**

[a] *, not analyzed in this study. Italic or boldface data represent nucleotide and amino acid sequences that are altered or not present, respectively, in the EDL933 strain used in this study (accession no. CP015855).

recent review, Vimr (16) postulated that chromosomally encoded NanS of *E. coli* may be responsible for the use of Neu5,9Ac$_2$ as a carbon source. The author demonstrated by *in silico* analyses that pathogenic *E. coli* bacteria carry numerous prophage-encoded ORFs which are homologous to the chromosomal *nanS* (*yjhS*) gene, which is part of the *nanCMS* operon of *E. coli* (16–18).

Landstorfer et al. (19) compared the transcriptome of strain EDL933 grown in LB medium to the transcriptome present after growth in different growth media such as minimal media, radish sprouts, or feces or under different conditions such as the use of differing pH values or antibiotics. Although transcription of the *nanS*-homologous ORFs was mostly downregulated under the tested conditions, it could be shown that the *nanS*-homologous ORFs are transcribed under certain conditions.

The major site of infection for EHEC is the large intestine, where carbon sources are limited. Mucus, which is composed of mucins containing various carbohydrates such as galactose, hexosamines, fucose, and neuraminic (sialic) acids, is thought to be one of the most important sources of energy supply. While highly fucosylated glycans were mainly found in the small intestine of human, the content of sialic acids increases along the colon (20). Following Neu5Ac, Neu5,9Ac$_2$ is the most frequently occurring neuraminic acid derivative in humans (16).

E. coli O157:H7 strain EDL933 was isolated from ground beef in 1982, during an EHEC outbreak in the United States (21). Since that time, EDL933 has been used as a reference EHEC strain in many laboratories worldwide. The genome sequence of EDL933 was first published in 2001 (22). Besides potential virulence factors and different metabolic pathways, 18 prophages were described. The genome of this strain has been resequenced by next-generation sequencing and has been published recently (23). The authors corrected the genome sequence by eliminating ambiguous base calls, closing a chromosomal gap, and revising the prophage regions. Therefore, a number of differences from the sequence published by Perna et al. (22) were found.

The aims of the current study were to identify, quantify, and analyze the *nanS*-homologous ORFs in the genome of EDL933 by *in silico* analyses, to clone selected *nanS*-homologous ORFs, and to determine their enzymatic activity. Generating isogenic and sequential deletions of the *nanS*-homologous ORFs demonstrated their importance for substrate utilization.

MATERIALS AND METHODS

Nomenclature of prophage-encoded *nanS*-homologous ORFs and use of a particular EDL933 derivative. The genome sequence of *E. coli* O157:H7 strain EDL933 was annotated twice (National Center of Biological Information [http://www.ncbi.nlm.nih.gov/] accession numbers AE005174.2 and NZ_CP008957.1). The experiments of this study were initially performed using the prior sequence published by Perna et al. (22) (AE005174.2). Therefore, many of the primers used in this study (see Table S1 in the supplemental material) were named according to the locus tags used in that sequence entry. Recent work was based on this sequence and the nomenclature that we used was also applied accordingly, e.g., gene locus z1466, plasmid pET-z1466-his (13–15).

Following the naming scheme used in the more recently reported sequence (NZ_CP008957.1), we renamed the corresponding loci. Due to their homologies to the chromosomal *nanS* gene and their prophage locations, we termed the prophage-encoded *nanS*-homologous genes the *nanS*-p genes and numbered them in accordance with the order in which they were deleted in our study (see below) (Table 1). The *nanS*-p8, *nanS*-p9, and *nanS*-p10 alleles published by Latif et al. (23) were not present in the genome of our EDL933 laboratory strain derivative (see below), and the numbering of these was determined according to their appearance in the genome sequence. The *nanS*-p1 allele is synonymous with the ORF formerly named z1466, but when materials from former work were used in this study, we kept this name.

Bioinformatic methods. The *nanS*-p alleles in the *E. coli* O157:H7 strain EDL933 genome were analyzed by BLAST analysis (24), and the prophage regions were aligned and depicted with ClustalW using Geneious R9 software (Biomatters Ltd., Auckland, New Zealand). In order to identify putative functional domains in the proteins, the NCBI BLASTP suite (http://blast.ncbi.nlm.nih.gov/Blast.cgi) was used. To analyze the phylogenetic relationships among NanS-p amino acid sequences in *E. coli* strain EDL933, a tree was calculated with the unweighted pair group method using average linkages (UPGMA) and Mega6 (25) from a Clustal alignment calculated with Bioedit, version 7.2.5, based on the amino acid sequences (26). The NanS sequence of EDL933 was used as an outgroup.

Bacterial strains, growth media, plasmids, and genome sequencing. Bacterial strains and plasmids used in this study are listed in Table 2. Plasmids were prepared with a QlAprep Spin Miniprep kit (Qiagen) according to the manufacturer's recommendation. Genomic DNA of EDL933 was prepared with a GenElute bacterial genomic DNA kit from Sigma-Aldrich according to the manufacturer's instructions and was sequenced on a PacBio RS2 device (Pacific Biosciences, Menlo Park, CA, USA) with a 10-kb size-selected insertion library and P6/C4 chemistry. One SMRT cell was used. *De novo* assembly (HGAP3 algorithm) was

TABLE 2 Bacterial strains and plasmids used in this study

Bacterial strain	Relevant characteristic(s)/plasmid	Reference or source
E. coli BL21(DE3)	F^- *ompT hsdS*$_B$(r_B^- m_B^-) *gal dcm* (DE3)	44
E. coli BL21(DE3)/pET-z1466-his		15
E. coli BL21(DE3)/pET-22b(+)	Donor for plasmid pET-22b(+) (Novagen, Merck)	15
E. coli BL21(DE3)/pET-*nanS*-p2-his		This study
E. coli BL21(DE3)/pET-*nanS*-p4-his		This study
E. coli JM 109	*recA1 endA1 gyrA96 thi hsdR17 supE44 relA1* λ^- Δ(*lac-proAB*) [F^- *traD36 proAB lacI*^{q} *lacZ*ΔM15]	45
E. coli BW25113/pKD46	Donor for plasmid pKD46	46
E. coli BW25141/pKD4	Donor for plasmid pKD4	46
E. coli BT340	*E. coli* DH5α carrying pCP20	47
E. coli C600	Wild type	48
E. coli C600Δ*nanS*	Deletion of *nanS*	This study
E. coli EDL933	Wild type	49
E. coli EDL933Δ*nanS*	Deletion of *nanS*	This study
E. coli EDL933Δ*nanS*Δ*nanS*-p1a	Deletion of *nanS*, *nanS*-p1a	This study
E. coli EDL933Δ*nanS*Δ*nanS*-p1a-p2	Deletion of *nanS*, *nanS*-p1a, *nanS*-p2	This study
E. coli EDL933Δ*nanS*Δ*nanS*-p1a-p3	Deletion of *nanS*, *nanS*-p1a, *nanS*-p2, *nanS*-p3	This study
E. coli EDL933Δ*nanS*Δ*nanS*-p1a-p4	Deletion of *nanS*, *nanS*-p1a, *nanS*-p2, *nanS*-p3, *nanS*-p4	This study
E. coli EDL933Δ*nanS*Δ*nanS*-p1a-p5	Deletion of *nanS*, *nanS*-p1a, *nanS*-p2, *nanS*-p3, *nanS*-p4, *nanS*-p5	This study
E. coli EDL933Δ*nanS*Δ*nanS*-p1a-p6	Deletion of *nanS*, *nanS*-p1a, *nanS*-p2, *nanS*-p3, *nanS*-p4, *nanS*-p5, *nanS*-p6	This study
E. coli EDL933Δ*nanS*Δ*nanS*-p1a-p7	Deletion of *nanS*, *nanS*-p1a, *nanS*-p2, *nanS*-p3, *nanS*-p4, *nanS*-p5, *nanS*-p6, *nanS*-p7	This study
E. coli EDL933Δ*nanS*Δ*nanS*-p6	Deletion of *nanS*, *nanS*-p6	This study

performed using SMRT Analysis version 2.3 software (Pacific Biosciences). HGAP3 settings were kept at the defaults, except for the expected genome size of 5 Mbp. LB medium was used for all routine purposes and consisted of 10 g/liter tryptone, 5 g/liter yeast extract, and 10 g/liter NaCl. The pH was adjusted to 7.0 with 1 N NaOH. The medium was autoclaved at 121°C for 15 min. Kanamycin sulfate (Sigma-Aldrich) or ampicillin sodium salt (Roth) was added at a final concentration of 50 μg/ml or 100 μg/ml, respectively, when required.

M9 minimal medium was prepared basically as described previously (27). Casamino Acids (Difco) were added to reach a final concentration of 0.1% (wt/vol), and glucose, Neu5,9Ac$_2$, or Neu5Ac was added from sterile stock solutions, when required, to reach a final concentration of 0.4% (wt/vol). For growth of *E. coli* C600 and derivatives, the medium was supplemented with leucine, threonine, and thiamine solutions at a final concentration of 0.0001% (wt/vol). SOC medium was prepared as described elsewhere (27).

Production of electrocompetent bacterial cells and electroporation. For transformation experiments, overnight cultures (LB medium, 37°C, 180 rpm) of *E. coli* BL21(DE3), *E. coli* JM109, *E. coli* C600, or EDL933 or the corresponding deletion mutants were diluted 1:100 in 50 ml SOC medium and the cultures were incubated at 37°C and 180 rpm up to an optical density at 600 nm of 0.5 to 0.7. If strains contained temperature-sensitive plasmids, they were grown at 30°C. The cultures were cooled on ice for 30 min, and the following steps were performed at 2°C: the cells were harvested by centrifugation at 4,000 × *g* for 8 min and washed first with 40 ml sterile ultrapure water (Millipore), then with 20 ml sterile ultrapure water, and finally with 2 ml sterile ultrapure water containing 10% (vol/vol) glycerol. The resulting cell pellets were dissolved in 100 μl sterile ultrapure water containing 10% glycerol. Aliquots (40 μl) of competent cells were used immediately for transformation or stored as electrocompetent cells at −70°C until use.

For transformation experiments, 40-μl volumes of electrocompetent cells were thawed on ice and mixed with 30 to 300 ng DNA in an ice-cold Gene Pulser/MicroPulser cuvette (Bio-Rad). The mixture was pulsed in a Gene Pulser Xcell electroporation system (Bio-Rad) at 25 μF, 200 Ω, and 2.5 kW for 5 ms. Following this, 1 ml of prewarmed SOC medium was added immediately. The suspension was then incubated for 1 h at 37°C (or at 30°C if temperature-sensitive plasmids were present) and then spread on LB agar with or without antibiotics.

Cloning of *nanS*-p alleles and recombinant production of NanS-p proteins. PCR products, which were used not only for confirmation but also for further steps, were purified using a Qiagen QlAquick PCR purification kit (Qiagen), and DNA concentrations were determined with a Nanodrop 2000 device (Thermo Scientific).

The *nanS*-p4 allele of *E. coli* O157:H7 strain EDL933 was amplified by PCR with primers z6054-NdeI-for and z6054-XhoI-rev (see Table S1 in the supplemental material).

The *nanS*-p2 allele showed high sequence similarity to *nanS*-p1. Therefore, it was amplified in a nested PCR approach using primers z3342-for and z3342-rev (see Table S1 in the supplemental material). Since *nanS*-p2 contains a NdeI restriction site, this site had to be removed to facilitate cloning. Therefore, the nucleotide sequence of the restriction site was modified by site-directed mutagenesis (28). Briefly, A-tailing with *Taq* polymerase modified the first-round nested PCR product of *nanS*-p2 according to the specifications of the manufacturer (New England BioLabs). It was then ligated to vector pGEM-T Easy (Promega) and transformed (see above) in *E. coli* JM109 according to the manufacturer's specifications. Correct insertion was verified by PCR with oligonucleotides M13-f and M13-r (see Table S1). A nucleotide exchange of adenine to cytosine was performed by PCR with primers z3342-18A/C-for and z3342-13T/G-rev at position 312 of *nanS*-p2, using Phusion High-Fidelity DNA polymerase (Thermo Scientific) (see Table S1). This procedure was modified from that described in a QuikChange II site-directed mutagenesis kit instruction manual (Agilent Technologies). The former NdeI restriction site changed to 5′-CCTATG-3′, which did not result in an amino acid change. A digestion with NdeI was performed subsequently (by the method described in a Thermo Scientific manual) to cut the unchanged template plasmid. The resulting target plasmid was used as the template for amplification of *nanS*-p2 with primers z3342-NdeI-for and z3342-XhoI-rev (see Table S1).

The PCR products of each approach were ligated into expression vector pET22b(+) as described elsewhere (17) and transformed into *E. coli* BL21(DE3) (see above). Correct insertion of the genes into pET22b(+) and the corresponding nucleotide sequences were confirmed by sequencing both DNA strands.

The resulting recombinant *E. coli* BL21(DE3)/pET-*nanS*-p2-his and BL21(DE3)/pET-*nanS*-p4-his strains, as well as the previously described BL21(DE3)/pET-z1466-his strain and the negative-control BL21(DE3)/

pET-22b(+) strain (15), were induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), harvested, and lysed, and the resulting protein solutions were purified and confirmed by Western blotting with an anti-His-tag antibody as described previously (15). The protein solutions were dialyzed in Slide-A-Lyzer dialysis cassettes (Thermo Scientific) (3.5-K-molecular-weight cutoff) in sodium-potassium buffer (10 mM Na_2HPO_4, 10 mM KH_2PO_4, pH 7.0) at 4°C in a 2.5-liter volume. The buffer was changed after 16 h, and the dialysis process continued for a further 4 h. The resulting protein preparations were designated NanS-p2-His and NanS-p4-His in addition to 933Wp42-His and the no-insertion control (NIC). The determination of protein concentrations by Bradford protein assay was carried out as previously described (15) using bovine serum albumin as a standard. The protein solutions were stored at −70°C.

Determination of the esterase activity of recombinant NanS-p proteins using Neu5,9Ac$_2$ and mucin. 5-*N*-Acetyl-9-*O*-acetyl neuraminic acid (Neu5,9Ac$_2$) was provided by Wolfgang Fessner and Ning He, Technical University Darmstadt, Darmstadt, Germany. It was synthesized as described previously (29) and analyzed by nuclear magnetic resonance (NMR) spectroscopy. The purity was >95%. 5-*N*-Acetyl neuraminic acid (Neu5Ac) was purchased from Applied BioTech, Austria. Neu5,9Ac$_2$ and Neu5Ac were used at a concentration of 6.25 mg/ml, corresponding to molar concentrations of 17.79 and 20.21 mmol, respectively, in 50 mM Tris-HCl (pH 8.0). Mucin from porcine stomach (Sigma-Aldrich) and mucin from bovine submaxillary gland (Merck Millipore), both dissolved in phosphate-buffered saline (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2HPO_4, pH 7.4), were used at a final concentration of 12.5 mg/ml. NanS-p2-His and NanS-p4-His as well as the 933Wp42-His positive control (15) and the NIC (see above) were diluted in 10 mM sodium potassium buffer (pH 7.0) to the concentrations used. The final protein concentrations used were 0.18 mg/liter for deacetylation of Neu5,9Ac$_2$ and 6.7 mg/liter for deacetylation of mucin. In this case, 1 katal (kat) esterase activity was defined as the release of 1 mol acetic acid per second.

Mixtures of 100 μl Neu5,9Ac$_2$ or the mucin solutions, respectively, were preincubated at 37°C. The reactions were started by adding 5 μl of the diluted protein solutions of the recombinant proteins or the positive control, respectively, with shaking at 300 rpm for 20 min. All assays were performed in triplicate. The reactions were stopped at 95°C and 500 rpm for 10 min. The released acetic acid concentrations were detected with a commercial kit (acetic acid; Boehringer Mannheim/R-Biopharm) according to the manufacturer's instructions. Neu5Ac and the NIC were used as negative controls.

The measurement range was 999 to 4,995 μmol acetic acid according to the instructions of the manufacturer. Amounts of substances below those values were assessed as representing negative conditions.

Biochemical characterization of recombinant NanS-p esterases. Esterase activity was determined by measuring hydrolysis of 4-methylumbelliferyl acetate (4-MUF-Ac) (Sigma-Aldrich) for investigation of the pH and temperature optima as described by Nübling et al. (15) with minor modifications as described below. The absorption measurements were carried out in an Infinite M200 microplate reader (Tecan). The 4-MUF-Ac stock solution was 5 mM. The buffers used had a molar concentration of 50 mM. Briefly, 80 μl of the corresponding recombinant esterase solution was preincubated for 4 min in 352 μl of the corresponding buffer system at the corresponding temperature. The reaction was started by adding 48 μl substrate stock solution and stopped by adding 100 μl 50% (vol/vol) acetic acid solution. A pH range of 4.0 to 9.0 in acetate buffer (pH 4.0 to 5.0), potassium phosphate buffer (pH 5.0 to 7.0), or Tris-HCl (pH 7.0 to 9.0), and a temperature range of 10 to 70°C were tested. The measurement of pH optimum was performed at 20°C or 40°C for NanS-p4-His or 933Wp42-His and NanS-p2-His, respectively. For detection of the temperature optima, a Tris-HCl (pH 7.0) buffer system was used. For the investigation of substrate specificity, 4-methylumbelliferyl caprylate (Santa Cruz Biotechnology), 4-methylumbelliferyl butyrate (Sigma-Aldrich), and 4-MUF-Ac were used as potential substrates. A commercial lipase (from *Candida rugosa*; Sigma-Aldrich) served as a positive control.

All measurements were performed in triplicate. Percentages of relative activity were calculated by assigning the highest level of enzymatic activity a value of 100%. For calculation of the enzymatic activity, determinations of calibration curves between 0 and 90 μmol 4-methylumbelliferone (Sigma-Aldrich) in each buffer system at each pH were performed. One katal (kat) of esterase$_{Muf\text{-}Ac}$ activity was defined as the release of 1 mol 4-methylumbelliferone within 1 s. For determinations of specific esterase$_{Muf\text{-}Ac}$ activity (measured as nanokatals per milligram), the activity (in nanokatals) was quantified in reference to 1 mg protein. The limit of quantification (LOQ) was calculated by multiplying the standard deviation of the blank values by a factor of 9.

Construction of *nanS*- and isogenic sequential *nanS*-p deletion mutants and growth experiments. Deletion mutagenesis was performed according to the method described by Datsenko and Wanner (28) using plasmids pKD46, pKD4, and pCP20 (Table 2). Primer pairs and PCR programs for deletion of *nanS* in *E. coli* C600 and EDL933 and sequential deletion of *nanS*-p1a to *nanS*-p7 in strain EDL933 are listed in Table S1 in the supplemental material. PCR products were purified and the DNA concentration was measured as described above. Deletions were confirmed by PCR with primer pairs surrounding the position where the deletion had taken place (see Table S1).

The resulting *E. coli* C600Δ*nanS* mutant strain was grown at an initial optical density at 600 nm of 0.1 in 5 ml M9 minimal medium with Neu5,9Ac$_2$ as a carbon source (see above) and a supplementation of 4 μg/ml NanS-p2-His, NanS-p4-His, or 933Wp42-His at 37°C and 180 rpm and compared to controls without supplement. Similarly, EDL933 and its mutant derivatives (Table 2) were cultivated with Neu5,9Ac$_2$, Neu5Ac, or glucose or without a carbohydrate source. The mutant strain EDL933Δ*nanS*Δ*nanS*-p1a-p7, with eight deletions, was cultivated with Neu5,9Ac$_2$ and with supplemental recombinant protein. Growth experiments were generally performed in triplicate with two measurements of optical density each per measurement point. Complementation experiments were performed in duplicate for the EDL933 deletion mutants.

Accession number(s). The sequence determined in this work was deposited in the NCBI database under accession no. CP015855 and organism designation *Escherichia coli* EDL933-1.

RESULTS

Bioinformatic analyses of the *nanS*-p alleles. The *E. coli* O157:H7 strain EDL933 genome as published by Latif et al. (23) contains a single chromosomal *nanS* gene and 10 *nanS*-p alleles (*nanS*-p1 to *nanS*-p10) (Fig. 1), the latter with high nucleotide sequence similarities to *nanS*. All *nanS*-p alleles in *E. coli* EDL933 are located in the genome of functional or cryptic lambdoid prophages. They are located between the respective late antiterminator Q gene and the lysis S gene, and the DNA region surrounding *nanS*-p is conserved in *E. coli* EDL933 (Fig. 2).

Sequence and phylogenetic analyses revealed that the corresponding NanS-p proteins of *E. coli* EDL933 cluster into three phylogroups (Fig. 1A and Table 1). NanS-p1 and NanS-p2, belonging to group I, are encoded by *nanS*-p1 and *nanS*-p2, located in close proximity to and downstream of stx_{2a} and stx_{1a} in the corresponding prophage genomes, and are 645 and 648 amino acids in length, respectively. They share overall amino acid similarity of 91.5%. Differences occur mainly between amino acids 34 and 78. NanS-p8 belongs to group II, with a size of 645 amino acids. The corresponding proteins of the seven *nanS*-p alleles *nanS*-p3 to *nanS*-p7, *nanS*-p9, and *nanS*-p10 belong to group III. In this group, the proteins consist of 616 or 617 amino acids and share sequence similarity of 95.0% to 99.7%.

The chromosomally encoded NanS is 326 amino acids in length

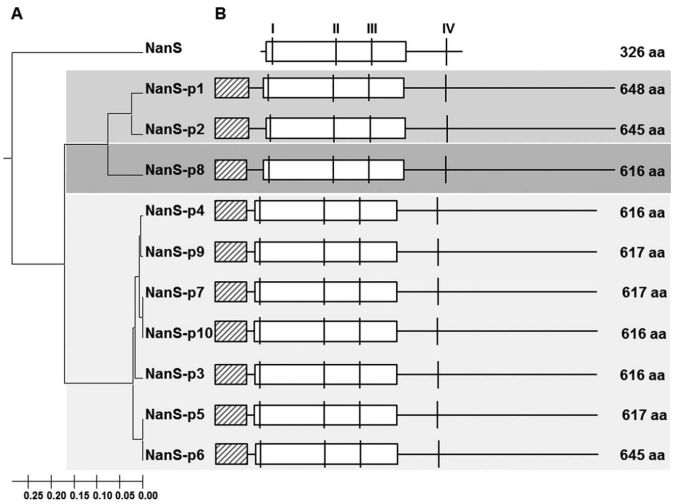

FIG 1 Phylogenetic analysis and structure of NanS and NanS-p protein sequences in *E. coli* O157:H7 strain EDL933. (A) The UPGMA tree of NanS and the NanS-p family was calculated with Mega6 (25) based on the amino acid (aa) sequences. (B) Structural organization of NanS and the NanS-p family. Potential functional domains (DUF, domain of unknown function) were predicted using BLAST. Horizontal boxes with black diagonal stripes depict DUF1737; white horizontal boxes indicate DUF303. Amino acid sequence blocks typical of the SGNH enzyme family of hydrolases as described previously (18) are depicted as vertical lines I, II, III, and IV. Thin horizontal continuous lines illustrate continuous amino acid chains.

and carries domain of unknown function 303 (DUF303), which has been previously linked to its esterase activity (18). In addition, NanS harbors four sequence blocks (blocks I to IV) (Fig. 1B) which have been described for the SGNH superfamily of enzymes (18). All NanS-p proteins also contain a DUF303 superfamily domain and possess those conserved sequence blocks which are present in enzymes of the SGNH superfamily. Moreover, the NanS-p proteins harbor a conserved DUF1737 domain of unknown function at their N-terminal ends. The DUF1737 sequence is similar to the CIII protein sequence of phage lambda from amino acid 14 to amino acid 37 which is the essential region for the activity of λCIII (30).

Since the genome sequence reported by Latif et al. (23) was generated with an EDL933 derivative deposited in the American Type Culture Collection (ATCC) and since we used a derivative that has been propagated over 30 years in German laboratories, we found some differences and decided to determine the genome

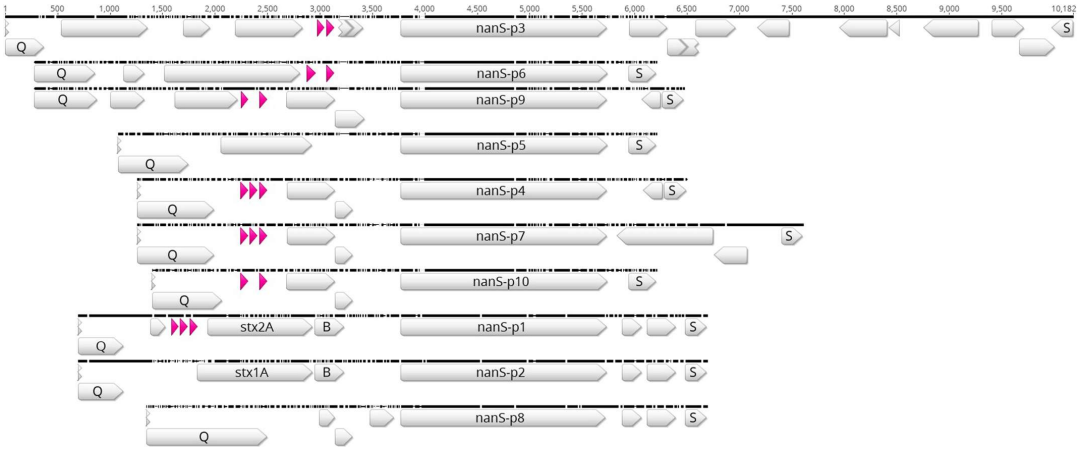

FIG 2 Comparison of the prophage regions between genes Q and S in *E. coli* O157:H7 strain EDL933. ClustalW alignment of prophage regions was performed using Geneious R9 software. Blank and pink arrows depict different prophage genes and transfer ribonucleic acids (tRNA), respectively. The compared prophage regions are flanked by an antiterminator Q gene and a lysis S gene. The stx_{1A} and stx_{1B} genes as well as the stx_{2A} and stx_{2B} genes are labeled.

sequence of our EDL933 derivative. A total of 63,172 reads with 651 Mbp were produced. The HGAP3 analysis produced a complete *de novo* genome sequence with 100-fold coverage over the entire molecules. Although we found the sequence to be mostly in accordance to the recently described EDL933 genome sequence (NZ_CP008957.1), there are some differences concerning the *nanS*-p alleles. Besides *nanS*, we discovered only 7 *nanS*-p alleles (*nanS*-p1 to *nanS*-p7) instead of the 10 present in the Latif genome. Alleles *nanS*-p8 to *nanS*-p10 were not present. The *nanS*-p1 allele (formerly z1466) which is located on the stx_{2a} prophage is slightly altered in its sequence in comparison with the z1466 sequence present in 933W. We designated it *nanS*-p1a. Alleles *nanS*-p2, *nanS*-p3, *nanS*-p4, *nanS*-p5, *nanS*-p6, and *nanS*-p7 showed the same nucleotide sequences as those previously published (23).

Functional characterization of recombinant NanS-p2 and NanS-p4 proteins. In order to characterize some NanS-p proteins in *E. coli* O157:H7 strain EDL933 enzymatically, we selected *nanS*-p2 and *nanS*-p4 for cloning. The *nanS*-p2 allele, which is located in the genome of the Stx1a-encoding phage (15), is highly homologous to the already described z1466 (*nanS*-p1) and belongs to group I. NanS-p4 belong to the more distantly related group III (Fig. 1). Since *nanS*-p8 was not present in our EDL933 laboratory strain, we did not analyze the group II protein. The selected *nanS*-p alleles were cloned into vector pET-22b(+), transformed into *E. coli* BL21(DE3), expressed as His-tagged proteins, and purified. The resulting proteins were designated NanS-p2-His and NanS-p4-His. For comparison purposes, 933Wp42-His (15) was used as a positive control in all experiments. Polyacrylamide gel electrophoresis of all preparations resulted in large (approximately 70-kDa) bands for all recombinant proteins. Western blotting with an anti-His-tag antibody demonstrated that all homologous proteins could be detected with high specificity (data not shown).

Characterization of the esterase activity of NanS-p2-His and NanS-p4-His. To demonstrate the esterase activity of NanS-p2-His and NanS-p4-His, cleavage of acetic acid from Neu5,9Ac_2 and from two different mucins was investigated with a commercial acetic acid detection assay. Release of acetic acid from Neu5,9Ac_2 demonstrated enzymatic activities of 7,578 ± 637 nkat/mg, 9,231 ± 300 nkat/mg, and 7,882 ± 294 nkat/mg for NanS-p2-His, NanS-p4-His, and the 933Wp42-His control, respectively, after incubation with the putative recombinant esterases. This result demonstrated slight differences in the esterase activities of the recombinant enzymes tested. Using the NIC, the free acetic acid molar concentration was below the recommended measurement range of the assay and was therefore considered a negative result. Moreover, acetic acid release from Neu5Ac was also considered a negative result. These results confirmed the assumption that NanS-p2-His and NanS-p4-His as well as 933Wp42-His (15) possess an *O*-acetyl esterase activity and cleave the acetyl group from C9 of Neu5,9Ac_2.

However, differences in the amounts of released acetic acid were observed when porcine stomach mucin and mucin from bovine submaxillary glands were used as substrates. Release of acetic acid from mucin from bovine submaxillary gland demonstrated specific enzymatic activities of 193 ± 12 nkat/mg, 196 ± 5 nkat/mg, and 200 ± 8 nkat/mg for NanS-p2-His, NanS-p4-His, and 933Wp42-His, respectively. Negative-control results were below the recommended measurement range.

In contrast, acetic acid release could not be detected after incubation of mucin from porcine stomach with the Neu5,9Ac_2 esterases. Kit controls were tested and gave positive results; thus, the possibility of any disturbing interaction between kit components and samples could be excluded.

Determination of characteristic enzymatic properties of the selected recombinant Neu5,9Ac_2 esterases. Due to the amounts of the different *nanS*-p alleles that were present in *E. coli* EDL933 and their proved Neu5,9Ac_2 esterase activity, we hypothesize that the different Neu5,9Ac_2 esterases of *E. coli* EDL933 may act under different environmental conditions. Initially, standard enzymatic tests were performed as described earlier (15). Using 4-methylumbelliferyl acetate, 4-methylumbelliferyl butyrate, and 4-methylumbelliferyl caprylate as synthetic esterase substrates, it was shown that NanS-p2-His, NanS-p4-His, and the positive control 933Wp42-His could cleave the acetic acid residues from 4-methylumbelliferyl acetate but not from 4-methylumbelliferyl butyrate and 4-methylumbelliferyl caprylate. Lipase activity could not be detected for any of the enzymes (data not shown).

In a next step, the temperature optima of the 3 Neu5,9Ac_2 esterases were determined in a range between 10°C to 70°C and the pH optimum was measured in a range between 4.0 and 9.0 in different buffer systems (Fig. 3). For NanS-p2-His and 933Wp42-His, a temperature of 50°C was determined to be optimal (Fig. 3A and B). Moreover, for NanS-p4-His, the temperature optimum was 40°C (Fig. 3C). The optimal pH range determined for all Neu5,9Ac_2 esterases in Tris-HCl buffer was between 7 and 9 (Fig. 3).

Supplementation of growth medium with recombinant Neu5,9Ac_2 esterases facilitates growth of *E. coli* C600Δ*nanS*. To test the hypothesis that the enzymatic activity of the prophage-encoded Neu5,9Ac_2 esterases allows growth of *E. coli* on Neu5,9Ac_2 as a carbon source, an *E. coli* C600 derivative lacking the chromosomal *nanS* gene was constructed and designated mutant strain C600Δ*nanS*. The wild-type *E. coli* C600 strain grew well in M9 minimal medium containing Neu5,9Ac_2 as a sole carbon source until an optical density at 600 nm of 2.1 was reached (Fig. 4).

In contrast, the C600Δ*nanS* deletion mutant grew only minimally in the medium (Fig. 4). Given that there should have been no growth, since *nanS* was deleted, the weak growth was probably due to the presence of a low concentration of Casamino Acids in the medium. Furthermore, the possibility cannot be excluded that there were side activities of other enzymes present which were produced from the strain. After addition of each of the Neu5,9Ac_2 esterases to the growth medium, the C600Δ*nanS* mutant regained the ability to grow on Neu5,9Ac_2. The growth rate and the obtained optical density were nearly identical to those seen with the *E. coli* C600 wild-type strain, suggesting that the Neu5,9Ac_2 esterases can complement the chromosomal NanS function.

Sequential deletion of *nanS*-p alleles induces loss of the ability of EDL933 to grown on Neu5,9Ac_2. Deletion mutagenesis in EDL933 was performed sequentially starting with the deletion of the chromosomal *nanS* gene. Subsequently, *nanS*-p1a, *nanS*-p2, *nanS*-p3, *nanS*-p4, *nanS*-p5, *nanS*-p6, and *nanS*-p7 were deleted in the same strain. This resulted in a mutant of EDL933, designated mutant strain EDL933Δ*nanS*Δ*nanS*-p1a-p7, which had eight deletions and which did not contain any *nanS*-p allele. Each deletion step was confirmed by PCR analysis. Furthermore, after

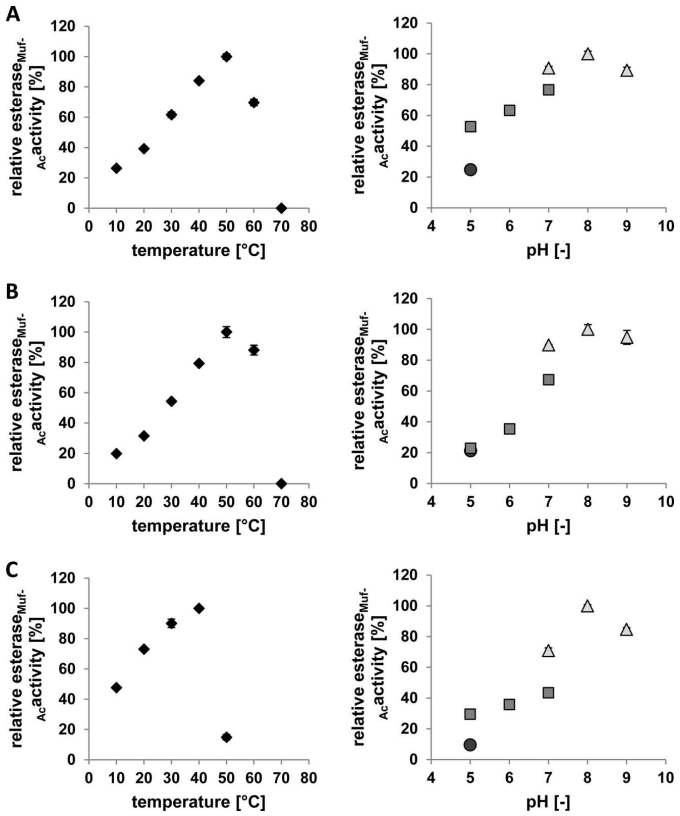

FIG 3 Measurement of temperature and pH optima of NanS-p2-His (A), 933Wp42-His (B), and NanS-p4-His (C). The substrate was 4-methylumbelliferyl acetate. Temperature conditions for measurement of pH optimum and buffer system (dot, acetate; square, potassium phosphate; triangle, Tris-HCl) were 40°C or 20°C for panels A and B or panel C, respectively. The temperature optimum was detected in Tris-HCl buffer (pH 7.0). In the case of temperature optima, 100% relative esterase$_{Muf-Ac}$ activity data represent 50.9 ± 0.9 nkat/mg (A), 35.3 ± 1.3 nkat/mg (B), and 53.1 ± 0.7 nkat/mg (C). In the case of pH optima, 100% relative esterase$_{Muf-Ac}$ activity data represent 41.5 ± 0.7 nkat/mg (A), 30.3 ± 0.9 nkat/mg (B), and 52.7 ± 0.4 nkat/mg (C).

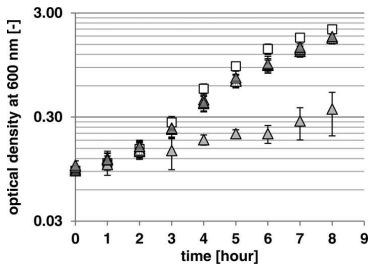

FIG 4 Supplementation of the growth medium of *E. coli* C600Δ*nanS* with the different recombinant NanS-p proteins. Data represent the results of growth of *E. coli* C600 (white squares) and the C600Δ*nanS* mutant (light gray triangles) on 0.4% Neu5,9Ac$_2$ in M9 minimal medium at 37°C and 180 rpm and of growth of the C600Δ*nanS* mutant with 4 μg/ml supplemented 933Wp42-His, NanS-p2-His, and NanS-p4-His (triangles with vertical stripes, dark gray triangles, and triangles with horizontal stripes, respectively).

deletion of all *nanS*-p alleles, we double-checked the number of homologs by PCR using primers which bind in all *nanS*-p alleles that are present in the EDL933 genome sequence (NZ_CP008957.1). We confirmed that there were no further *nanS*-p alleles in the EDL933Δ*nanS*Δ*nanS*-p1a-p7 mutant strain.

Preliminary tests did not show noteworthy growth defects on Neu5,9Ac$_2$ of mutant strains EDL933Δ*nanS*Δ*nanS*-p1a-p2, EDL933Δ*nanS*Δ*nanS*-p1a-p3, EDL933Δ*nanS*Δ*nanS*-p1a-p4, and EDL933Δ*nanS*Δ*nanS*-p1a-p5 compared to the wild type (data not shown). Therefore, we performed an experiment in which we grew mutant strains EDL933, EDL933Δ*nanS*, EDL933Δ*nanS*Δ*nanS*-p1a-p5, EDL933Δ*nanS*Δ*nanS*-p1a-p6, and EDL933Δ*nanS*Δ*nanS*-p1a-p7 in M9 minimal medium with Neu5,9Ac$_2$ as a carbon source (Fig. 5A). Mutant strains EDL933Δ*nanS* and EDL933Δ*nanS*Δ*nanS*-p1a-p5 grew just as well as the EDL933 wild-type strain. Interestingly, starting with the 7-fold deletion mutant, EDL933Δ*nanS*Δ*nanS*-p1a-p6, the strain lost its ability to grow on Neu5,9Ac$_2$, although *nanS*-p7 was still present. To exclude any undesired defects in the *nan* operons

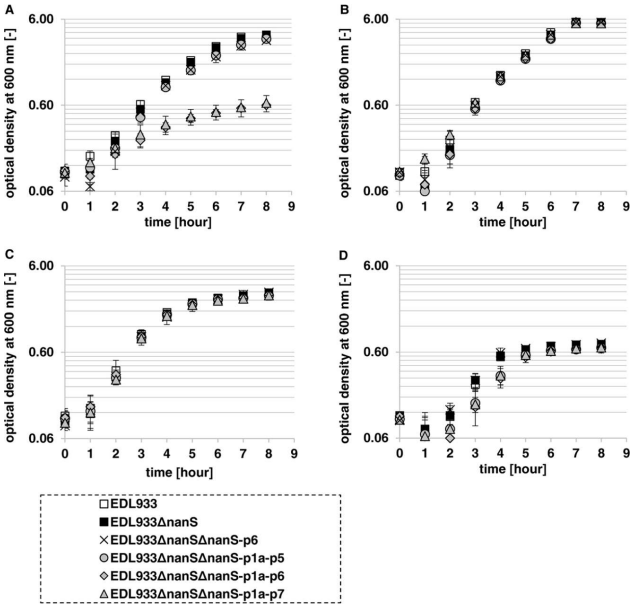

FIG 5 Growth curves of *E. coli* O157:H7 strain EDL933 and selected EDL933 deletion mutants with different numbers of deleted *nanS*-p alleles. Data represent the results of cultivation in M9 minimal medium with 0.4% Neu5,9Ac$_2$ (A), Neu5Ac (B), or glucose (C) or without a special added carbohydrate source (D) at 37°C and 180 rpm over 8 h. (A and D) The EDL933 wild-type strain and the EDL933Δ*nanS*, EDL933Δ*nanS*Δ*nanS*-p6, and EDL933Δ*nanS*Δ*nanS*-p1a-p5 mutant strains grew well to approximately the same optical density (A), while the EDL933Δ*nanS*Δ*nanS*-p1a-p6 and EDL933Δ*nanS*Δ*nanS*-p1a-p7 mutant strains grew just up to the basal optical density (A and D). (B and C) All mutants grew as well as the wild-type strain on the positive controls with Neu5Ac (B) or glucose (C) as the carbon source. (D) Basal growth was measured in the negative control, where no carbon source had been added.

or in growth ability in general, we let this strain and the other mutant strains grow on Neu5Ac and glucose in parallel (Fig. 5B and C). To test whether the growth defect of mutant strain EDL933Δ*nanS*Δ*nanS*-p1a-p6 was attributable to the deletion of *nanS*-p6 or whether it was a consequence of the number of *nanS*-p alleles that had been deleted, we constructed a double mutant with deletion of *nanS* and *nanS*-p6 (mutant strain EDL933Δ*nanS*Δ*nanS*-p6) and let it grow as before. Indeed, mutant strain EDL933Δ*nanS*Δ*nanS*-p6 grew well (Fig. 5A), and we reasoned that the growth defect of mutant strain EDL933Δ*nanS*Δ*nanS*-p1a-p6 on Neu5,9Ac$_2$ was attributable to the quantity of deletions and thus that a gene dose effect exists. Mutant strain EDL933Δ*nanS*Δ*nanS*-p1a-p7 could not grow on Neu5,9Ac$_2$, as expected, and a basal level of growth could be measured in all cases (Fig. 5D) that was attributable to the supplemented Casamino Acids.

In order to analyze whether the ability of mutant strain EDL933Δ*nanS*Δ*nanS*-p1a-p7 to grow could be restored, M9 minimal medium was supplemented with recombinant 933Wp42-His, NanS-p2-His, or NanS-p4-His. It was found that the growth of the complemented strain could be restored to approximately the wild-type level in all three cases (Fig. 6).

DISCUSSION

E. coli O157:H7 strain EDL933 has been used as an EHEC reference strain worldwide for more than 30 years. Its genome sequence was analyzed and published twice (NCBI accession numbers AE005174.2 and NZ_CP008957.1), with the most notable sequence differences located in the repeated prophage regions (22, 23). In this context, Vimr (16) described 7 complete and 5 partial *nanS*-homologous alleles in the EDL933 genome, while we found

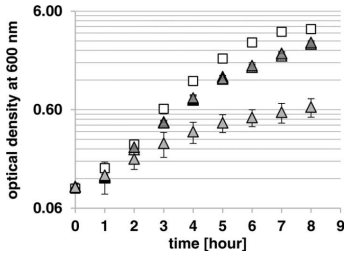

FIG 6 Supplementation of the growth medium of the EDL933Δ*nanS*Δ*nanS*-p1a-p7 mutant with different recombinant NanS-p proteins. Data represent the results of growth of the EDL933 wild-type strain (white squares) and the EDL933Δ*nanS*Δ*nanS*-p1a-p7 mutant strain (light gray triangles) on 0.4% Neu5,9Ac$_2$ in M9 minimal medium at 37°C and 180 rpm and of growth of the EDL933Δ*nanS*Δ*nanS*-p1a-p7 mutant strain with 4 μg/ml supplemented 933Wp42-His, NanS-p2-His, and NanS-p4-His (triangles with vertical stripes, dark gray triangles, and triangles with horizontal stripes, respectively).

10 complete *nanS*-p alleles using the recently determined nucleotide sequence (23). Moreover, analyzing our EDL933 laboratory strain by whole-genome sequencing, only seven *nanS*-p alleles were detected. We hypothesize, in accordance with a recent publication (31), that the EDL933 genomes of different isolates were different from that of the original isolate due to evolutionary events that took place over the period of time since the strain was first isolated (21). Different storage and inoculation techniques could probably trigger such mutations. Such events have been demonstrated earlier by Papadopoulos et al. (32), who have shown the dynamic nature of the *E. coli* genome. Moreover, the specific genetic structures of the phages encoding NanS-p6 and NanS-p9, localized in tandem and separated only by transposases, would probably lead to instability in this region (data not shown).

In silico analysis of EDL933 identified the location of all *nanS*-p alleles in the region of late gene expression of prophages. Moreover, the positions of *nanS*-p1 and *nanS*-p2 are directly downstream of the positions of stx_{2a} and stx_{1a}, respectively, in the genome of Stx phages. In fact, *nanS*-p1 and stx_{2a} or *nanS*-p2 *and* stx_{1a} are cotranscribed after phage induction (13). The other *nanS*-p alleles are located in the genome of non-Stx phages but positioned in the corresponding region. Detailed investigation within the *nanS*-p alleles showed differences from the *nanS* gene. The additional nucleotide sequences which are present downstream of the typical sequence blocks for the SGNH enzyme family may be relevant for a further enzymatic function or for structural purposes, but the details are currently not known. A further region (DUF1737) which shows similarities to a prophage repressor gene was identified at the N terminus, but its function is currently unknown (see Fig. 1).

The presence of multiple *nanS*-p alleles in the EHEC genomes raises the issue of the functionalities and activities of the corresponding gene products. Based on the bioinformatic analyses and molecular experiments performed during this study, there is strong body of evidence that derivatives of Neu5Ac may serve as substrates for carbon nutrition in the large intestine. The presence of multiple *nanS*-p genes suggests that their gene products may give an advantage for growth under certain environmental conditions. By performing deletion experiments in EDL933, we were able to show that it did not lose the ability to grow on Neu5,9Ac_2 if at least 2 *nanS*-p alleles were still present. That indicated that large mutations or other genetic events could happen without loss of carbohydrate utilization in the large intestine.

To evade naturally occurring colonization resistance in the gut, EHEC bacteria have to outcompete the gut microbiome for at least one nutrient (33). Furthermore, *E. coli* O157:H7 needs mechanisms to avoid host defense and to penetrate the mucus layer to reach the gut epithelium (34). Human gastrointestinal mucus consists of two layers, a thick loosely mucus layer and a thinner layer attached to the mucosa (35). While the inner layer is nearly sterile, the outer mucus layer forms a microbial niche that also includes nonmucolytic bacteria (36, 37). One of those bacterial species is *E. coli* (38). *E. coli* bacteria lack the ability to degrade the complex sugar chains present in the MUC2 mucus glycoprotein. However, they are actually in equilibrium with the turnover rate of mucus. One possibility is explained in the "Restaurant" hypothesis, which implies that *E. coli* bacteria grow in synergy with gut anaerobes which can degrade large polysaccharide chains to mono- or disaccharides (34). Indeed, it was shown that commensal *E. coli* strain MG1655 as well as *E. coli* O157:H7 strain EDL933 can grow well on mucus-derived saccharides such as *N*-acetylglucosamine, gluconate, galactose, *N*-acetyl neuraminic acids, and others (39). Interestingly, the same authors showed that deletion defects in one of the *nan* operons, responsible for growth on *N*-acetyl neuraminic acid, led to a growth defect in *E. coli* MG1655 but not in EDL933.

Regarding the distribution of sialic acids in the human gut, the monoacetylated Neu5Ac and the diacetylated Neu5,9Ac_2 are represented in large amounts (16). In the mucin, they play a role in protection of the core protein against microbes. Nevertheless, there are some anaerobes which can remove the sialic acids from the glycoprotein by sialidases (40–42). Mostly, sialidases are specialized to cleave only a few bonds of sialic acids to the glycoprotein or to have narrow substrate specificity. Thus, *Tannerella forsythia*, a pathogenic member of the phylum *Bacteroidetes*, uses NanS to deacetylate glycolytically bound Neu5,9Ac_2 and Neu5Gc,9Ac (43). The residual Neu5Ac or Neu5Gc can then be separated from the glycoprotein with NanH, which is the sialidase corresponding to *Bacteroidetes*, and the neuraminic acids can then be used as the carbon source. Although *E. coli* does not express any sialidases, it still possesses the *nanS* gene (17, 18).

We hypothesize that the multiple Neu5,9Ac_2 esterases could provide *E. coli* O157:H7 with a colonization advantage by enabling it to outcompete the commensal *E. coli* strains in the gut for Neu5,9Ac_2.

While NanS is located in the periplasm of the bacterial cells (16), it is probable that, similarly to Stx, additional Neu5,9Ac_2 esterases may be released by phage lysis into the environment. This hypothesis is justified by the localization of the genes in the late regulated prophage regions. Therefore, we added different Neu5,9Ac_2 esterases to the medium containing Neu5,9Ac_2 and grew laboratory *E. coli* strain C600Δ*nanS* and EDL933 mutant strain EDL933Δ*nanS*Δ*nanS*-p1a-p7 under this set of conditions. As expected, both grew to almost the same optical density as the wild-type strain without deletions in the *nanS* gene and the *nanS*-p gene, respectively.

The different Neu5,9Ac_2 esterases could be part of a larger network of substrate utilization enzymes in the gut. This could be a reason for the high number of homologous genes with similar functions present in EHEC strains.

ACKNOWLEDGMENTS

We thank Ning He and Wolfgang Fessner, University of Darmstadt, Germany, for providing 5-*N*-acetyl-9-*O*-acetyl neuraminic acid. We also thank Melanie Schneider and Markus Kranz for skillful technical assistance.

FUNDING INFORMATION

This work, including the efforts of Herbert Schmidt, was funded by Deutsche Forschungsgemeinschaft (DFG) (Schm1360/6-1).

REFERENCES

1. **Nataro JP, Kaper JB.** 1998. Diarrheagenic *Escherichia coli*. Clin Microbiol Rev **11**:142–201.
2. **Karmali MA, Petric M, Lim C, Fleming PC, Arbus GS, Lior H.** 1985. The association between idiopathic hemolytic uremic syndrome and infection by verotoxin-producing *Escherichia coli*. J Infect Dis **151**:775–782. http://dx.doi.org/10.1093/infdis/151.5.775.
3. **Bielaszewska M, Dobrindt U, Gartner J, Gallitz I, Hacker J, Karch H, Muller D, Schubert S, Alexander SM, Sorsa LJ, Zdziarski J.** 2007. Aspects of genome plasticity in pathogenic *Escherichia coli*. Int J Med

Microbiol **297**:625–639. http://dx.doi.org/10.1016/j.ijmm.2007.03.001.
4. **Wirth T, Falush D, Lan R, Colles F, Mensa P, Wieler LH, Karch H, Reeves PR, Maiden MC, Ochman H, Achtman M.** 2006. Sex and virulence in *Escherichia coli*: an evolutionary perspective. Mol Microbiol **60**:1136–1151. http://dx.doi.org/10.1111/j.1365-2958.2006.05172.x.
5. **Nguyen Y, Sperandio V.** 2012. Enterohemorrhagic *E. coli* (EHEC) pathogenesis. Front Cell Infect Microbiol **2**:90. http://dx.doi.org/10.3389/fcimb.2012.00090.
6. **Kaper JB, Nataro JP, Mobley HL.** 2004. Pathogenic *Escherichia coli*. Nat Rev Microbiol **2**:123–140. http://dx.doi.org/10.1038/nrmicro818.
7. **Paton AW, Srimanote P, Talbot UM, Wang H, Paton JC.** 2004. A new family of potent AB(5) cytotoxins produced by Shiga toxigenic *Escherichia coli*. J Exp Med **200**:35–46. http://dx.doi.org/10.1084/jem.20040392.
8. **Karch H, Denamur E, Dobrindt U, Finlay BB, Hengge R, Johannes L, Ron EZ, Tonjum T, Sansonetti PJ, Vicente M.** 2012. The enemy within us: lessons from the 2011 European *Escherichia coli* O104:H4 outbreak. EMBO Mol Med **4**:841–848. http://dx.doi.org/10.1002/emmm.201201662.
9. **Unkmeir A, Schmidt H.** 2000. Structural analysis of phage-borne stx genes and their flanking sequences in Shiga toxin-producing *Escherichia coli* and *Shigella dysenteriae* type 1 strains. Infect Immun **68**:4856–4864. http://dx.doi.org/10.1128/IAI.68.9.4856-4864.2000.
10. **Fogg PC, Saunders JR, McCarthy AJ, Allison HE.** 2012. Cumulative effect of prophage burden on Shiga toxin production in *Escherichia coli*. Microbiology **158**:488–497. http://dx.doi.org/10.1099/mic.0.054981-0.
11. **Wagner PL, Waldor MK.** 2002. Bacteriophage control of bacterial virulence. Infect Immun **70**:3985–3993. http://dx.doi.org/10.1128/IAI.70.8.3985-3993.2002.
12. **Schmidt H.** 2001. Shiga-toxin-converting bacteriophages. Res Microbiol **152**:687–695. http://dx.doi.org/10.1016/S0923-2508(01)01249-9.
13. **Herold S, Siebert J, Huber A, Schmidt H.** 2005. Global expression of prophage genes in *Escherichia coli* O157:H7 strain EDL933 in response to norfloxacin. Antimicrob Agents Chemother **49**:931–944. http://dx.doi.org/10.1128/AAC.49.3.931-944.2005.
14. **Polzin S, Huber C, Eylert E, Elsenhans I, Eisenreich W, Schmidt H.** 2013. Growth media simulating ileal and colonic environments affect the intracellular proteome and carbon fluxes of enterohemorrhagic *Escherichia coli* O157:H7 strain EDL933. Appl Environ Microbiol **79**:3703–3715. http://dx.doi.org/10.1128/AEM.00062-13.
15. **Nübling S, Eisele T, Stober H, Funk J, Polzin S, Fischer L, Schmidt H.** 2014. Bacteriophage 933W encodes a functional esterase downstream of the Shiga toxin 2a operon. Int J Med Microbiol **304**:269–274. http://dx.doi.org/10.1016/j.ijmm.2013.10.008.
16. **Vimr ER.** 2013. Unified theory of bacterial sialometabolism: how and why bacteria metabolize host sialic acids. ISRN Microbiol **2013**:816713. http://dx.doi.org/10.1155/2013/816713.
17. **Steenbergen SM, Jirik JL, Vimr ER.** 2009. YjhS (NanS) is required for *Escherichia coli* to grow on 9-*O*-acetylated *N*-acetylneuraminic acid. J Bacteriol **191**:7134–7139. http://dx.doi.org/10.1128/JB.01000-09.
18. **Rangarajan ES, Ruane KM, Proteau A, Schrag JD, Valladares R, Gonzalez CF, Gilbert M, Yakunin AF, Cygler M.** 2011. Structural and enzymatic characterization of NanS (YjhS), a 9-*O*-acetyl *N*-acetylneuraminic acid esterase from *Escherichia coli* O157:H7. Protein Sci **20**:1208–1219. http://dx.doi.org/10.1002/pro.649.
19. **Landstorfer R, Simon S, Schober S, Keim D, Scherer S, Neuhaus K.** 2014. Comparison of strand-specific transcriptomes of enterohemorrhagic *Escherichia coli* O157:H7 EDL933 (EHEC) under eleven different environmental conditions including radish sprouts and cattle feces. BMC Genomics **15**:353. http://dx.doi.org/10.1186/1471-2164-15-353.
20. **Robbe C, Capon C, Coddeville B, Michalski JC.** 2004. Structural diversity and specific distribution of *O*-glycans in normal human mucins along the intestinal tract. Biochem J **384**:307–316. http://dx.doi.org/10.1042/BJ20040605.
21. **Riley LW, Remis RS, Helgerson SD, McGee HB, Wells JG, Davis BR, Hebert RJ, Olcott ES, Johnson LM, Hargrett NT, Blake PA, Cohen ML.** 1983. Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. N Engl J Med **308**:681–685. http://dx.doi.org/10.1056/NEJM198303243081203.
22. **Perna NT, Plunkett G, III, Burland V, Mau B, Glasner JD, Rose DJ, Mayhew GF, Evans PS, Gregor J, Kirkpatrick HA, Posfai G, Hackett J, Klink S, Boutin A, Shao Y, Miller L, Grotbeck EJ, Davis NW, Lim A, Dimalanta ET, Potamousis KD, Apodaca J, Anantharaman TS, Lin J, Yen G, Schwartz DC, Welch RA, Blattner FR.** 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. Nature **409**:529–533. http://dx.doi.org/10.1038/35054089.
23. **Latif H, Li HJ, Charusanti P, Palsson BO, Aziz RK.** 14 August 2014. A gapless, unambiguous genome sequence of the enterohemorrhagic *Escherichia coli* O157:H7 strain EDL933. Genome Announc http://dx.doi.org/10.1128/genomeA.00821-14.
24. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic local alignment search tool. J Mol Biol **215**:403–410. http://dx.doi.org/10.1016/S0022-2836(05)80360-2.
25. **Tamura K, Stecher G, Peterson D, Filipski A, Kumar S.** 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol **30**:2725–2729. http://dx.doi.org/10.1093/molbev/mst197.
26. **Hall TA.** 2016. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/N. Nucleic Acids Symp Ser **41**:95–98.
27. **Maniatis T, Sambrook J, Fritsch EF.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
28. **Hutchison CA, III, Phillips S, Edgell MH, Gillam S, Jahnke P, Smith M.** 1978. Mutagenesis at a specific position in a DNA sequence. J Biol Chem **253**:6551–6560.
29. **Ogura H, Furuhata K, Sato S, Anazawa K, Itoh M, Shitori Y.** 1987. Synthesis of 9-*O*-acyl- and 4-*O*-acetyl-sialic acids. Carbohydr Res **167**:77–86. http://dx.doi.org/10.1016/0008-6215(87)80269-0.
30. **Kobiler O, Rokney A, Oppenheim AB.** 2007. Phage lambda CIII: a protease inhibitor regulating the lysis-lysogeny decision. PLoS One **2**:e363. http://dx.doi.org/10.1371/journal.pone.0000363.
31. **Fellner L, Huptas C, Simon S, Muhlig A, Scherer S, Neuhaus K.** 7 April 2016. Draft genome sequences of three European laboratory derivatives from enterohemorrhagic *Escherichia coli* O157:H7 strain EDL933, including two plasmids. Genome Announc http://dx.doi.org/10.1128/genomeA.01331-15.
32. **Papadopoulos D, Schneider D, Meier-Eiss J, Arber W, Lenski RE, Blot M.** 1999. Genomic evolution during a 10,000-generation experiment with bacteria. Proc Natl Acad Sci U S A **96**:3807–3812. http://dx.doi.org/10.1073/pnas.96.7.3807.
33. **Freter R, Brickner H, Botney M, Cleven D, Aranki A.** 1983. Mechanisms that control bacterial populations in continuous-flow culture models of mouse large intestinal flora. Infect Immun **39**:676–685.
34. **Conway T, Cohen PS.** June 2015. Commensal and pathogenic *Escherichia coli* metabolism in the gut. Microbiol Spectr http://dx.doi.org/10.1128/microbiolspec.MBP-0006-2014.
35. **Atuma C, Strugala V, Allen A, Holm L.** 2001. The adherent gastrointestinal mucus gel layer: thickness and physical state in vivo. Am J Physiol Gastrointest. Liver Physiol **280**:G922–G929.
36. **Li H, Limenitakis JP, Fuhrer T, Geuking MB, Lawson MA, Wyss M, Brugiroux S, Keller I, Macpherson JA, Rupp S, Stolp B, Stein JV, Stecher B, Sauer U, McCoy KD, Macpherson AJ.** 2015. The outer mucus layer hosts a distinct intestinal microbial niche. Nat Commun **6**:8292. http://dx.doi.org/10.1038/ncomms9292.
37. **Johansson ME, Phillipson M, Petersson J, Velcich A, Holm L, Hansson GC.** 2008. The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. Proc Natl Acad Sci U S A **105**:15064–15069. http://dx.doi.org/10.1073/pnas.0803124105.
38. **Poulsen LK, Lan F, Kristensen CS, Hobolth P, Molin S, Krogfelt KA.** 1994. Spatial distribution of *Escherichia coli* in the mouse large intestine inferred from rRNA in situ hybridization. Infect Immun **62**:5191–5194.
39. **Fabich AJ, Jones SA, Chowdhury FZ, Cernosek A, Anderson A, Smalley D, McHargue JW, Hightower GA, Smith JT, Autieri SM, Leatham MP, Lins JJ, Allen RL, Laux DC, Cohen PS, Conway T.** 2008. Comparison of carbon nutrition for pathogenic and commensal *Escherichia coli* strains in the mouse intestine. Infect Immun **76**:1143–1152. http://dx.doi.org/10.1128/IAI.01386-07.
40. **Tailford LE, Owen CD, Walshaw J, Crost EH, Hardy-Goddard J, Le GG, de Vos WM, Taylor GL, Juge N.** 2015. Discovery of intramolecular trans-sialidases in human gut microbiota suggests novel mechanisms of mucosal adaptation. Nat Commun **6**:7624. http://dx.doi.org/10.1038/ncomms8624.
41. **Crost EH, Tailford LE, Le GG, Fons M, Henrissat B, Juge N.** 2013. Utilisation of mucin glycans by the human gut symbiont *Ruminococcus*

gnavus is strain-dependent. PLoS One **8:**e76341. http://dx.doi.org/10.1371/journal.pone.0076341.

42. **Nees S, Veh RW, Schauer R.** 1975. Purification and characterization of neuraminidase from *Clostridium perfringens*. Hoppe Seylers Z Physiol Chem **356:**1027–1042. http://dx.doi.org/10.1515/bchm2.1975.356.s1.1027.
43. **Phansopa C, Kozak RP, Liew LP, Frey AM, Farmilo T, Parker JL, Kelly DJ, Emery RJ, Thomson RI, Royle L, Gardner RA, Spencer DI, Stafford GP.** 2015. Characterization of a sialate-O-acetylesterase (NanS) from the oral pathogen *Tannerella forsythia* that enhances sialic acid release by NanH, its cognate sialidase. Biochem J **472:**157–167. http://dx.doi.org/10.1042/BJ20150388.
44. **Studier FW, Moffatt BA.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J Mol Biol **189:**113–130. http://dx.doi.org/10.1016/0022-2836(86)90385-2.
45. **Yanisch-Perron C, Vieira J, Messing J.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119. http://dx.doi.org/10.1016/0378-1119(85)90120-9.
46. **Datsenko KA, Wanner BL.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A **97:**6640–6645. http://dx.doi.org/10.1073/pnas.120163297.
47. **Cherepanov PP, Wackernagel W.** 1995. 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. http://dx.doi.org/10.1016/0378-1119(95)00193-A.
48. **Miller JH.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
49. **O'Brien AO, Lively TA, Chen ME, Rothman SW, Formal SB.** 1983. *Escherichia coli* O157:H7 strains associated with haemorrhagic colitis in the United States produce a *Shigella dysenteriae* 1 (SHIGA) like cytotoxin. Lancet **i:**702.

Kapitel 3

Growth advantage of *Escherichia coli* O104:H4 strains on 5-*N*-acetyl-9-*O*-acetyl neuraminic acid as a carbon source is dependent on heterogeneous phage-borne NanS-p esterases

Nadja Saile[1], Lisa Schwarz[1], Kristina Eißenberger[1], Jochen Klumpp[2], Florian W. Fricke[3], Herbert Schmidt[1]

[1] Department of Food Microbiology and Hygiene, Institute of Food Science and Biotechnology, University of Hohenheim, Stuttgart, Germany

[2] Department of Health Sciences and Technology, Institute of Food, Nutrition and Health, ETH Zurich, Zurich, Switzerland

[3] Department of Nutrigenomics, Institute of Biological Chemistry and Nutritional Science, University of Hohenheim, Stuttgart, Germany

Publiziert in International Journal of Medical Microbiology, Band 308, Seiten 459-468.

DOI: 10.1016/j.ijmm.2018.03.006. Mit freundlicher Genehmigung von Elsevier.

International Journal of Medical Microbiology 308 (2018) 459–468

Contents lists available at ScienceDirect

International Journal of Medical Microbiology

journal homepage: www.elsevier.com/locate/ijmm

Growth advantage of *Escherichia coli* O104:H4 strains on 5-*N*-acetyl-9-*O*-acetyl neuraminic acid as a carbon source is dependent on heterogeneous phage-Borne nanS-p esterases

Nadja Saile[a], Lisa Schwarz[a], Kristina Eißenberger[a], Jochen Klumpp[c], Florian W. Fricke[b], Herbert Schmidt[a,*]

[a] Department of Food Microbiology and Hygiene, Institute of Food Science and Biotechnology, Garbenstrasse 28, University of Hohenheim, 70599 Stuttgart, Germany
[b] Department of Nutrigenomics, Institute of Biological Chemistry and Nutritional Science, Garbenstrasse 30, University of Hohenheim, 70599 Stuttgart, Germany
[c] Department of Health Sciences and Technology, Institute of Food, Nutrition and Health, Schmelzbergstrasse 7, ETH Zurich, 8092 Zurich, Switzerland

ARTICLE INFO

Keywords:
E. coli O104:H4
LB226692
C227-11Φcu
nanS-p alleles
Mucin
Neu5,9Ac$_2$

ABSTRACT

Enterohemorrhagic *E. coli* (EHEC) are serious bacterial pathogens which are able to cause a hemorrhagic colitis or the life-threatening hemolytic-uremic syndrome (HUS) in humans. EHEC strains can carry different numbers of phage-borne *nanS*-p alleles that are responsible for acetic acid release from mucin from bovine submaxillary gland and 5-*N*-acetyl-9-*O*-acetyl neuraminic acid (Neu5,9Ac$_2$), a carbohydrate present in mucin. Thus, Neu5,9Ac$_2$ can be transformed to 5-*N*-acetyl neuraminic acid, an energy source used by *E. coli* strains. We hypothesize that these NanS-p proteins are involved in competitive growth of EHEC in the gastrointestinal tract of humans and animals. The aim of the current study was to demonstrate and characterize the *nanS-p* alleles of the 2011 *E. coli* O104:H4 outbreak strain LB226692 and analyze whether the presence of multiple *nanS*-p alleles in the LB226692 genome causes a competitive growth advantage over a commensal *E. coli* strain.

We detected and characterized five heterogeneous phage-borne *nanS*-p alleles in the genome of *E. coli* O104:H4 outbreak strain LB226692 by *in silico* analysis of its genome. Furthermore, successive deletion of all *nanS*-p alleles, subsequent complementation with recombinant NanS-p13-His, and *in vitro* co-culturing experiments with the commensal *E. coli* strain AMC 198 were conducted. We could show that *nanS*-p genes of *E. coli* O104:H4 are responsible for growth inhibition of strain AMC 198, when Neu5,9Ac$_2$ was used as sole carbon source in co-culture. The results of this study let us suggest that multiple *nanS*-p alleles may confer a growth advantage by outcompeting other *E. coli* strains in Neu5,9Ac$_2$ rich environments, such as mucus in animal and human gut.

1. Introduction

Between May and July 2011, a large outbreak of diarrhea and the hemolytic-uremic syndrome (HUS) was recorded in Germany and Europe, which affected especially middle-aged women after consumption of contaminated fenugreek sprouts (Frank et al., 2011; King et al., 2012; Robert Koch-Institut, 2012). The outbreak was caused by a particular enterohemorrhagic *E. coli* strain of serotype O104:H4. Before that outbreak, only sporadic cases of *E. coli* O104:H4 infections had been described (Bae et al., 2006; Mellmann et al., 2008; Rauw et al., 2014; Scheutz et al., 2011).

Classical "EHEC" can be described as a group of pathogenic *E. coli* that produces one or more phage-encoded Shiga toxins (Stx) and contain the pathogenicity island "locus of enterocyte effacement (LEE)" (Kaper, 1998). The O104:H4 outbreak strain produced Stx2a and lacked the LEE (Bielaszewska et al., 2011; Brzuszkiewicz et al., 2011). Moreover, its genetical backbone was found to be closely related to those of enteroaggregative *E. coli* (EAEC), with an additional *stx* $_{2a}$-containing prophage in the chromosomal integration site *wrbA* (Brzuszkiewicz et al., 2011; Mellmann et al., 2011; Rasko et al., 2011). The genome of the O104:H4 outbreak strain showed high DNA sequence identity with the Stx-negative EAEC strain 55989, which was isolated from an HIV-positive adult in Africa in 1990 (Brzuszkiewicz et al., 2011; Mossoro et al., 2002). In culture, the strain was able to build the EAEC-typical aggregative stacked-brick pattern on HEp-2 cells (Bielaszewska et al., 2011; Tietze et al., 2015). Therefore, Brzuszkiewicz et al. (2011) described it as a new pathotype, namely entero-aggregative-haemorrhagic *Escherichia coli* (EAHEC). Further virulence factors, that have been

* Corresponding author.
E-mail address: herbert.schmidt@uni-hohenheim.de (H. Schmidt).

https://doi.org/10.1016/j.ijmm.2018.03.006
Received 31 January 2018; Received in revised form 15 March 2018; Accepted 19 March 2018

Table 1
E. coli strains and plasmids used in this study.

E. coli strains	Relevant characteristic	Reference
LB226692	serotype O104:H4	Mellmann et al. (2011)
C227-11ϕcu	serotype O104:H4, cured from the stx_{2a}- phage	Zangari et al. (2013)
C227-11ϕcuΔ*nanS*-p11	serotype O104:H4, deletion of *nanS*-p11	This study
C227-11ϕcuΔ*nanS*-p14	serotype O104:H4, deletion of *nanS*-p14	This study
C227-11ϕcuΔ*nanS*-p11,12	serotype O104:H4, deletion of *nanS*-p11 and *nanS*-p12	This study
C227-11ϕcuΔ*nanS*-p11,12,13	serotype O104:H4, deletion of *nanS*-p11, *nanS*-p12 and *nanS*-p13	This study
C227-11ϕcuΔ*nanS*-p11,12,13,14	serotype O104:H4, deletion of *nanS*-p11, *nanS*-p12, *nanS*-p13 and *nanS*-p14	This study
AMC 198		German Collection of Microorganisms and Cell Cultures (DSMZ) (DSM No. 787) (ATCC 11229)
BW25113/pKD46	Donor for plasmid pKD46	Datsenko and Wanner (2000)
BW25141/pKD4	Donor for plasmid pKD4	Datsenko and Wanner (2000)
BT340	*E. coli* DH5α carrying plasmid pCP20	Cherepanov and Wackernagel (1995)
XL1-Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F´*proAB lacI*^{q} *Z*Δ*M15* Tn*10* (Tet^r)]	Stratagene
XL1-Blue/pKEC1.5	Donor for pKEC1.5 (derivative of plasmid pKD46 containing a chloramphenicol resistance gene instead of the original ampicillin resistance gene)	This study
DH5α/pK18	Donor for plasmid pK18	Pridmore (1987)
BL21(DE3)	F–*ompT hsdSB*(rB −,mB −) *gal dcm* (DE3)	Studier and Moffatt (1986)
BL21(DE3)/pET22b(+)	BL21 (DE3) carring plasmid pET-22b(+) (Novagen, Merck)	Nübling et al. (2014)
BL21(DE3)/pET-*nanS*-p13	pET-22b(+), *nanS*-p13	This study

found in the outbreak strain, are the IrgA homologue adhesin encoding gene (*iha*), the aggregative adherence fimbriae type I (AAF/I), the *rpoS* gene, encoding a regulator for stress response, including high acid tolerance, the occurrence of plasmid-encoded TEM-1 and CTX-M-15 beta-lactamases, responsible for the ESBL phenotype, and the existence of an unusual combination of the three serine protease autotransporters of *Enterobacteriaceae* (SPATE) genes, *sepA*, *sigA* and *pic* (Brzuszkiewicz et al., 2011; Mellmann et al., 2011; Rasko et al., 2011). SPATE proteins seem to be involved in mucosal colonization and damage (Boisen et al., 2009; Dautin, 2010; Harrington et al., 2009).

Infections with EAEC belong to the most frequent *E. coli*-associated illnesses occurring in developing countries especially concerning children, immunosuppressed individuals or travellers (Harrington et al., 2005; Mossoro et al., 2002). The reservoir of typical EAEC (expressing the AggR regulon), such as the 2011 outbreak strain, seems to be the human gastrointestinal tract, different from STEC or atypical EAEC (not expressing the AggR regulon) that are preferentially found in ruminants (Uber et al., 2006; Wieler et al., 2011). Therefore, it is expected that typical EAEC are well adapted to humans and persist in the human gut microbiota.

Recently, we reported that EHEC O157:H7 strain EDL933 express several prophage-encoded 5-*N*-acetyl-9-*O*-acetyl neuraminic acid esterases (NanS-p), which are able to cleave acetyl residues from bovine submaxillary gland mucin and 5-*N*-acetyl-9-*O*-acetyl neuraminic acid (Neu5,9Ac$_2$) (Nübling et al., 2014; Saile et al., 2016). As a result of this cleavage, Neu5,9Ac$_2$ is metabolized to *N*-acetyl neuraminic acid (Neu5Ac). Neu5Ac inactivates the transcriptional repressor NanR of *E. coli* and therefore induces the *nan*-operons *nanATEK-yhcH*, *nanCMS* and *yjhBC* generating *N*-acetylmannosamine and pyruvate, which can be further degraded by the tricarboxylic acid cycle and/or glycolysis (Vimr, 2013).

The molecular structure of NanS-p proteins are deposited by domains of unknown functions 1737 and 303 (SASA) and comprise amino acid sequence blocks typical of the SGNH enzyme family of hydrolases (Rangarajan et al., 2011; Rangel et al., 2016; Saile et al., 2016). A special feature of the *nanS*-p genes in *E. coli* O157:H7 is their localization in prophages between the antiterminator Q and the lysis gene S, which is in contrast to the chromosomally located *nanS* of *E. coli* (as part of the *nanCMS* operon), that encodes also a Neu5,9Ac$_2$-esterase (Rangarajan et al., 2011; Steenbergen et al., 2009). Independently from each other, Rangel et al. (2016) and our group could recently state, that NanS-p can compensate the function of NanS in the Neu5,9Ac$_2$ catabolism in *E. coli* O157:H7 strain EDL933 (Rangel et al., 2016; Saile et al., 2016). But in contrast to *nanS* the *nanS*-p genes are not under the control of NanR (Rangel et al., 2016). Different from the situation in *nanS*, the *nanS*-p genes are frequently found in pathogenic *E. coli*, and in particular in clinical relevant EHEC serotypes (Vimr, 2013). In a former study of Herold et al. (2005), we could show that several *nanS*-p alleles were transcribed after norfloxacin induction. The data of this study confirmed also the cotransduction of *nanS*-p1 (z1466) and the Shiga toxin 2a gene together with the late phase prophage genes in O157:H7 strain EDL933 (Herold et al., 2005). In a later study, Polzin et al. (2013) demonstrated by differential 2D gel electrophoresis that NanS-p proteins (933Wp42) were 40-fold overexpressed in simulated colonic environmental medium (SCEM) under aerobic conditions in strain EDL933. These data show already evidence that *nanS*-p genes are functional and can be expressed under various conditions, especially under such conditions, where the SOS response is induced.

It is our hypothesis, that EHEC could have an advantage towards commensal *E. coli* and other *Enterobacteriaceae* with respect to substrate utilization in the human colon due to the presence of multiple *nanS*-p genes (Saile et al., 2016).

In the current study, we demonstrate that the O104:H4 outbreak strain LB226692 carries five prophage-located *nanS*-p genes. We characterized their location, analyzed their heterogeneous sequences and found that the presence of the *nanS*-p genes leads to a growth advantage on Neu5,9Ac$_2$ compared to the commensal *E. coli* strain AMC 198. The presence of multiple functional *nanS*-p alleles therefore increase the virulence of some pathogenic *E. coli* strains by providing growth advantages in the gut and facilitating colonization.

2. Materials and methods

2.1. E. coli strains and plasmids

E. coli strains used in this study are listed in Table 1. Strains were routinely grown in LB-broth (10 g/l tryptone, 5 g/l yeast extract, and 10 g/l NaCl). The pH was adjusted to 7.0 with 1 N NaOH. The antibiotics ampicillin (Roth), kanamycin (Sigma-Aldrich) or chloramphenicol (Roth) were added in final concentrations of 100, 50 or 25 µg/mL, respectively, when needed. Plasmids were prepared from overnight cultures of the respective *E. coli* strain using the QlAprep Spin Miniprep Kit (Qiagen).

2.2. Whole genome sequencing and in silico analysis

Genomic DNA was isolated using the GenElute Bacterial Genomic DNA Kit (Sigma-Aldrich) from 3 mL of an overnight culture of *E. coli* O104:H4 strain LB226692 (Table 1), analyzed with a Nanodrop 2000 device (Thermo Scientific) and sequenced on a PacBio system at the ETH Zurich, Switzerland, as previously described (Saile et al., 2016). Gene finding and functional annotation was performed with CloVR (Angiuoli et al., 2011; http://clovr.org/). The sequence of the chromosome was used for *in silico* analysis in this study.

The presence and the localisation of the *nan*-operons *nanATEK-yjhH*, *nanCMS* and *yjhBC* was identified by mapping *nanATEK-yjhH*, *nanCMS* and *yjhBC* from *E. coli* strain K-12 substrain MG1655 (NCBI Reference Sequence: NC_000913.3) to the LB226692 sequence with Geneious software version 9.1.8 (Biomatters Ltd., Auckland, New Zealand). The number, nucleotide sequences and genome positions of the *nanS*-p alleles in strain LB226692, were also identified with Geneious by mapping *nanS*-p alleles of *E. coli* O157:H7 strain EDL933 [NZ_CP008957.1; (Latif et al., 2014)] to the LB226692 sequence.

Prophage regions were identified using the PHAge Search Tool Enhanced Release (PHASTER) tool (Arndt et al., 2016; Zhou et al., 2011). The position of *nanS*-p in the corresponding prophage sequences was determined and depicted using the ClustalW Alignment contained in the Geneious software. NanS-p amino acid sequences were aligned equally to identify the sequence identities among the different NanS-p proteins. Functional and unknown domains in NanS-p amino acid sequences were identified by the HMMER tool (Finn et al., 2015).

For analysis of the distribution of *nanS*-p and the presence of *nan*-operons in selected strains, the following chromosomal DNA sequences were used: O104:H4 strains 2009EL-2050 (NC_018650.1), 2009EL-2071 (NC_018661.1), 55989 (NC_011748.1) and O44:H18 strain 042 (NC_017626.1). The *nanS*-p alleles were sequentially enumerated (from *nanS*-p16 to *nanS*-p30). The identification was performed by blastn analysis (Anon, 2017) and a phylogenetic tree was conducted using the UPGMA tree build method with the genetic distance model Jukes-Cantor and Geneious Tree Builder also including the NanS-p proteins from LB226692 and EDL933.

For identification of signal peptides, we used SecretomeP 2.0 (Anon, 2018a) and SignalP 4.1 (Anon, 2018b) software tools (Bendtsen et al., 2005; Petersen et al., 2011). SignalP predict signal peptide cleavage sites responsible for the Sec-dependent secretion pathway. SecretomeP identifies proteins that are putatively secreted in a Sec-independent way.

2.3. Cloning of the sialidase gene BT_0455 of Bacteroides thetaiotaomicron and nanS-p13 from E. coli O104:H4

Bacteroides thetaiotaomicron VPI 5482 (DSM No. 2079, Leibniz-Institute DSMZ- German collection of microorganisms and cell cultures) was incubated at 37 °C on Columbia agar with 5% sheep blood (Becton, Dickinson and Company) in an anaerobic jar using Anaerocult A (Merck Millipore) to establish an anaerobic atmosphere. Genomic DNA was prepared as described above. Using the NCBI reference sequence NC_004663.1 (Xu et al., 2003), the open reading frame of the gene BT_0455 from *B. thetaiotaomicron,* which encodes the sialidase BTSA (Park et al., 2013) was amplified without its putative signal sequence using primers BT0455-*Nde*I-f2 (5′-CCCCATATGTCAGACACCGTTTTTG TAC-3′) and BT0455-*Xho*I-r (5′-AAACTCGAGTCGAATCAAATCTTT CAG-3′).

The *nanS*-p13 allele was amplified with primers *nanS*-p13-NdeI-f (5′-CCCCATATGGCATTTAAACACTATGATGTTGTCAGGG-3′) and *nanS*-p13-XhoI-r (5′-AAACTCGAGTGCGGCTGCTCCAGC-3′). Genomic DNA of strain LB226692 was used as a template.

The PCR products were restricted with *Nde*I and *Xho*I (Thermo Scientific) and cloned in expression vector pET22b(+), restricted with the same enzymes, and transformed into *E. coli* BL21(DE3). Correct insertion and correctness of the sequence were confirmed by DNA sequencing as previously described (Nübling et al., 2014)

2.4. Expression and puri cation of BTSA-His and NanS-p13-His

The recombinant proteins were expressed and purified using Ni-NTA-beads as described previously (Nübling et al., 2014), except dialysis of BTSA-His was performed against 50 mM Tris-HCl pH 7.0 in a Slide-A-Lyzer dialysis cassette with a molecular-weight cutoff of 3.5 K (Thermo Scientific). After purification and buffer exchange, the protein concentrations and purities were determined by a Bradford assay and SDS-PAGE, respectively (Nübling et al., 2014).

2.5. Detection of enzymatic activity of NanS-p13-His

The esterase activity of NanS-p13-His was measured using a 5 mM 4-methylumbelliferyl acetate (4-MUF-Ac) (Sigma-Aldrich) solution as substrate. The temperature and pH optima in different buffer systems (acetate, potassium-phosphate, Tris-HCl) were determined as recently described (Saile et al., 2016). Briefly, the recombinant NanS-p13-His protein was diluted 1:50 in 10 mM sodium potassium buffer (pH 7.0). The temperature optimum was determined in Tris-HCl (pH 7.0) and the pH optimum was analyzed at 40 °C. The enzymatic activity katal (kat) was defined as the release of 1 mol 4-methylumbelliferone within 1 s. The highest specific esterase activity (kat/ mg protein) was set as 100% and the other values were calculated accordingly.

2.6. Construction of deletion mutant strains

Deletion mutants were constructed using the lambda red recombinase system (Datsenko and Wanner, 2000). Since *E. coli* O104:H4 strain LB226692 is resistant against ampicillin, the *bla* gene of plasmid pKD46, encoding a TEM-1 beta-lactamase was replaced by a *cat* gene encoding a chloramphenicol-acetyltransferase. The *cat* gene was amplified using plasmid pCP20 as a template with primers P-cat PvuI for (5′-ATACGATCGAGCGCTGATGTCCGGC-3′) and cat-PvuI rev (5′-ATAC GATCGTTACGCCCCGCCCTGCCA-3′). With these primers, a PCR product of 1032 bp was generated that contains the *cat* gene and a 375 bp fragment upstream of the *cat* open reading frame. Both, the PCR product and pKD46 were digested with restriction enzyme *Pvu*I, then ligated together using T4 DNA ligase (Thermo Scientific) over 16 h at 4 °C, transformed in chemical competent *E. coli* XL1-Blue (Inoue et al., 1990) by a heat shock at 42 °C for 90 s and regenerated in SOC medium (Maniatis et al., 1982) for 1 h. The construct was confirmed by DNA sequencing (see above). The resulting plasmid was designated pKEC1.5.

Electrocompetent cells for transformation were generated as recently described (Saile et al., 2016). If necessary, 25 µg/mL chloramphenicol was used instead of ampicillin. Transformation efficiency of cells was determined using a defined concentration of plasmid pK18.

Primers for deletion and the confirmation of deletion were constructed using the LB226692 genome sequence. The presence of *nanS*-p11, *nanS*-p12, *nanS*-p13 and *nanS*-p14 as well as the absence of *nanS*-p15 and *stx*2a in C227-11ϕcu was confirmed either using primers binding in the surrounding sequences of the gene (nanSp11f, nanSp11r; O104nanSP2-f, O104nanSP2-r; O104nanSp3-f, O104nanSP3-r; O104nanSP4-f, O104nanSP4-r) or, for *nanS*-p15, with primers binding in the gene (nanSp15f, nanSp15r) (Table 2). As positive control, genomic DNA of LB226692 was used.

Using the plasmids pKEC1.5, pKD4 and pCP20, the deletion primers described in Table 2, and the corresponding PCR programs (Table 2), isogenic single deletion mutants C227-11ϕcuΔnanS-p11 and C227-11ϕcuΔ*nanS*-p14, the double, triple and fourfold deletion mutants C227-11ϕcuΔ*nanS*-p11,12, C227-11ϕcuΔ*nanS*-p11,12,13, and C227-11ϕcuΔ*nanS*-p11,12,13,14 (Table 1) were constructed. The success of the deletion experiments was confirmed by PCR, using the primers described above, which are complementary to the respective gene

Table 2
Primers and PCR programs used in this study. Primers were constructed during this study, if not otherwise indicated.

Primer designation	Primer sequence (5′ 3′)	Target sequence	PCR program
nanSp11f	TATCCGACAGTAATCACTCTGCG	*nanS*-p11 flanking regions	300 s, 95 °C; 30 x (30 s, 95 °C; 30 s, 56 °C; 180 s, 68 °C); 300 s, 68 °C
nanSp11r	TCATTTCAATCTCAGCTGCTACC		
O104nanSP2-f	CACCAGACGGGAGTAACACA	*nanS*-p12 flanking regions	300 s, 95 °C; 30 x (30 s, 95 °C; 30 s, 54 °C; 120 s, 68 °C); 300 s, 68 °C
O104nanSP2-r	CTCGCAGACTGTTTGTCTTAAG		
O104nanSP3-f	GGCGACTGGTGACGGTATAATCTCTGATTATC	*nanS*-p13 flanking regions	300 s, 95 °C; 30 x (30 s, 95 °C; 30 s, 60 °C; 150 s, 68 °C); 300 s, 68 °C
O104nanSP3-r	CTTCCATTCCTGCCGCTGG		
O104nanSP4-f	CATTTTTCAGATGGTGGAAAGAACC	*nanS*-p14 flanking regions	300 s, 95 °C; 30 x (30 s, 95 °C; 30 s, 55 °C; 180 s, 68 °C); 300 s, 68 °C
O104nanSP4-r	GTCCTGGATGAAACCAGTAAGC		
nanSp15f	GGGCATCATCGACAGTATCAGAG	*nanS*-p15	300 s, 95 °C; 30 x (30 s, 95 °C; 30 s, 57 °C; 45 s, 68 °C); 300 s, 68 °C
nanSp15r	TGTTCACTTCCACCGCAATG		
non-stx-for	GCAATCGGTCACTGGTTCG	*stx*2a flanking regions	300 s, 95 °C; 30 x (30 s, 95 °C; 30 s, 50 °C; 90 s, 68 °C); 300 s, 68 °C
non-stx-rev2	AGATTACACTTGTTACCCAC		
Del-Z1466-for[2]	ATGGCATTTAAACACTATGATGTTGTCAGGGCGGCGTCGCCGgcgattgtgtaggctggagc [a,b]	pKD4 to generate *nanS*-p11/p14-Kan^R cassette	300 s, 95 °C; 30 x (30 s, 95 °C; 30 s, 59 °C; 120 s, 68 °C); 300 s, 68 °C
D-O104nanSP1/4-r	TTATGATGATGCCGGTGCATTTATCTCCAGCACCAGACTTTCcatggtccatatgaatatcctcc [a]		
D-O104nanSP2-f	ATGACCAGTAAGAACTTTGCACTGATTACAGCCATGACACAGgcgattgtgtaggctggagc [a]	pKD4 to generate *nanS*-p12-Kan^R cassette	
D-O104nanSP2-r	TTATGCTGCGGGATTATTGATTTCCAGCACCAGGCTGCCGATcatggtccatatgaatatcctcc [a]		
Del-Z1466-for[2]	ATGGCATTTAAACACTATGATGTTGTCAGGGCGGCGTCGCCGgcgattgtgtaggctggagc [a,b]	pKD4 to generate *nanS*-p13-Kan^R cassette	
D-O104nanSP3-r	TTATGCGGCTGCTCCAGCTTTGTTTGCTTTAACTTCCACCGTcatggtccatatgaatatcctcc [a]		

[a] lower case signals the complementary part to pKD4 and upper case signals the complementary part to the gene, that should be deleted.
[b] Saile et al. (2016).

flanking regions (Table 2).

2.7. Growth experiments

All growth experiments were performed in duplicate or in triplicate in M9 minimal medium (Maniatis et al., 1982), supplemented with 0.4% (w/v) glucose (Merck Millipore), 0.4% (w/v) Neu5,9Ac $_2$ (provided by Wolfgang Fessner TU Darmstadt), (Ogura et al., 1987), Neu5Ac (Applied Biotech) or with 0.8% (w/v) mucin derived from bovine submaxillary gland (Merck Millipore). Overnight cultures were grown in M9 minimal medium with glucose. The initial optical density at 600 nm (OD_{600}) was adjusted to 0.1. Growth competition experiments were performed by inoculation of fresh medium with equal amounts of the two strains to an initial OD_{600} of 0.2.

For growth control, strains C227-11ΦcuΔ*nanS*-p11, C227-11ΦcuΔ*nanS*-p14, C227-11ΦcuΔ*nanS*-p11,12, C227-11ΦcuΔ*nanS*-p11,12,13, C227-11ΦcuΔ*nanS*-p11,12,13,14 and strain AMC 198 were grown using glucose as a carbon source, and strains C227-11Φcu and C227-11ΦcuΔ*nanS*-p11,12,13,14 additionally on Neu5Ac to monitor putative physiological changes resulting from deletion mutagenesis. All mutant strains constructed in this study, C227-11Φcu and AMC 198 were grown in monoculture with Neu5,9Ac $_2$ as a sole carbon source. For growth of the control strains C227-11ΦcuΔ*nanS*-p11,12,13 and C227-11ΦcuΔ*nanS*-p11,12,13,14, M9 medium was supplemented with recombinant NanS-p13-His to a final concentration of 4 µg/mL. Since from our former study it became clear that there are no major differences in enzyme activity of the different *nanS*-p alleles with respect to complementation of the mutants, we have chosen NanS-p13-His as a representative (Saile et al., 2016). For competition experiments, strain AMC 198 together with C227-11Φcu, C227-11ΦcuΔ*nanS*-p11,12,13 or C227-11ΦcuΔ*nanS*-p11,12,13,14 were grown in parallel in M9 minimal medium with Neu5,9Ac $_2$. Furthermore, strains C227-11Φcu and AMC 198 were grown in co-culture with mucin from bovine submaxillary gland mucin as a carbon source with our without a supplement of 3.86 µg/mL BTSA-His. Equally, C227-11ΦcuΔ*nanS*-p11,12,13,14 and AMC 198 were co-cultured.

Bacterial growth was determined by measuring the OD_{600} and/or the viable counts every two hours. Therefore, serial dilutions in 0.9% (w/v) sodium chloride were plated in duplicate on LB agar and/or LB containing 50 µg/mL streptomycin (Sigma-Aldrich). The viable counts for strain AMC 198 in competition experiments with C227-11ϕcu or mutant strains were calculated by subtracting colony counts on LB agar and LB agar with streptomycin, because strain AMC 198 is susceptible to streptomycin, and O104:H4 is not. The generation time is the time interval that is needed to double the bacterial cell count. It was calculated by subtracting the time of the logarithmic phase of the bacterial growth curve with the count of cell divisions within this time. Bacterial growth curves were analyzed by paired Student's *t* test using Microsoft Excel (version 15.33), if necessary.

2.8. Accession number

The genomic DNA sequence of *E. coli* O104:H4 strain LB226692, determined in this study, was deposited in the NCBI database under the accession numbers CP024992-CP024994.

3. Results

3.1. In silico analysis of genes involved in sialic acid catabolism

E. coli O104:H4 strain LB226692 was isolated during the large outbreak on May 24, 2011 from a stool sample of a HUS patient (Mellmann et al., 2011) and was obtained from the German National HUS reference centre, Prof. Helge Karch, Münster, Germany. Since only a draft genome sequence of LB226692 with multiple contigs was available from the NCBI databases, a de-novo assembled genome sequence was generated using single molecule realtime (SMRT) sequencing, resulting in three individual contigs for the chromosome of 5,299,162 bp and the two plasmids of 106,829 and 94,453 bp length. While the chromosomal sequence was the basis for the bioinformatic analysis, the plasmid DNA was not analyzed in this study.

In a first genome analysis with PHASTER, 12 prophage-associated

Table 3
PHASTER analysis of prophage regions in LB226692 chromosome. Score signals to which group, incomplete (score < 70), questionable (score 70–90) or intact (score > 90), the prophage region belongs.

Region	Region Length	Completeness	Score	Total Proteins	Region Position	Most Common Phage	GC%	*nanS*-p occurrence
1	26.1 kb	incomplete	40	13	347727–373848	PHAGE_Entero_phi_NC_023693	47.12	n.d.
2	28.6 kb	intact	120	22	982129–1010728	PHAGE_Entero_BP_4795_NC_004813	51.53	n.d.
3	24.3 kb	intact	120	20	1569809–1594207	PHAGE_Salmon_SJ46_NC_031129	57.81	n.d.
4	27.7 kb	incomplete	30	33	3079685–3107406	PHAGE_Shigel_Sf6_NC_005344	45.74	n.d.
5	8.7 kb	incomplete	20	16	3477404–3486185	PHAGE_Klebsi_vB_KpnP_KpV475_NC_031025	50.87	n.d.
6	63.6 kb	intact	150	71	3648411–3712040	PHAGE_Entero_mEp460_NC_019716	51.41	*nanS*-p11
7	46.7 kb	intact	150	54	3852409–3899135	PHAGE_Entero_P88_NC_026014	52.17	*nanS*-p12
8	66.8 kb	questionable	80	69	4028392–4095211	PHAGE_Shigel_POCJ13_NC_025434	49.78	*nanS*-p13
9	47.3 kb	intact	150	60	4396677–4444030	PHAGE_Entero_lambda_NC_001416	51.48	*nanS*-p14
10	21.7 kb	incomplete	20	28	4644514–4666231	PHAGE_Erwini_vB_EamM_Asesino_NC_031107	51.52	n.d.
11	62.8 kb	intact	120	69	4758173–4821035	PHAGE_Escher_P13374_NC_018846	50.26	*nanS*-p15
12	46.3 kb	intact	150	59	5082037–5128388	PHAGE_Entero_lambda_NC_001416	50.69	n.d.

n.d. = not detected.

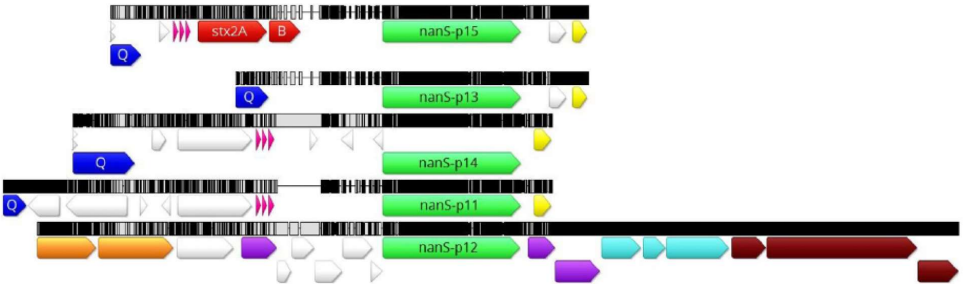

Fig. 1. ClustalW alignment of *nanS*-p (green) alleles and flanking regions of prophages present in strain LB226692. The genes *nanS*-p11, *nanS*-p13, *nanS*-p14 and *nanS*-p15 (downstream of *stx* genes (red)) are located between the antiterminator Q (dark blue) and the lysis gene S (yellow). The n*anS*-p12 allele is located between the phage structural genes (capsid = orange, tail = violet, baseplate = turquoise, fiber = dark red). Pink and white colored arrows depict tRNA genes and other open reading frames, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

DNA-regions were identified (Table 3). In 5 prophages, *nanS*-p alleles were identified, that were designated *nanS*-p11, −12, −13, −14, and −15 (Table 3) in accordance with the proposed nomenclature (Saile et al., 2016). The *nanS*-p15 allele is located downstream of stx_{2a}, in the respective phage. More detailed analysis demonstrated similar genetic localizations for *nanS*-p11, *nanS*-p13, *nanS*-p14 and *nanS*-15, i.e. in the late transcribed regions of their corresponding prophages, between the genes Q and S (Fig. 1). In contrast, *nanS*-p12 is localized in the late phage region encoding capsid and tail proteins (Fig. 1).

The open reading frames (ORF) of the *nanS*-p11, *nanS*-p12, *nanS*-p13, *nanS*-p14 and *nanS*-p15 alleles have sizes of 1854, 1881, 1911, 1854 and 1938 bp, respectively. The putative amino acid sequences demonstrated that they also belong to the SGNH hydrolases superfamily. More accurately, as first defined by Rangarajan et al. (2011), they belong to the SGNH type II class which showed differences to the common SGNH hydrolase type I class in their conserved sequence blocks (Fig. 2). For example, the common asparagine in block III is absent. Besides the functional esterase regions (SASA), they also carry the domain of unknown function (DUF) 1737 as described for the EDL933 *nanS*-p alleles (Saile et al., 2016) (Fig. 2). The amino acid identities between the LB226692 NanS-p sequences are between 48 and 88% (data not shown).

By using SignalP 4.1, we could not identify any N-terminal signal peptides (data not shown). However, the NanS-p proteins were grouped to non-classically secreted protein with SecretomeP 2.0 software. The software generate a score between 0 and 1 whereas a score above 0.5 is an indication for secreted proteins. NanS-p11, NanS-p12, NanS-p13, NanS-p14 and NanS-p15 receive scores of 0.947, 0.951, 0.940, 0.948 and 0.947, respectively.

We also investigated the LB226692 sequence for the presence of the chromosomal NanR-regulated *nan*-operons and compared that to other stx_{2a}-positive *E. coli* O104:H4 strains and to the stx_{2a}-negative, but closely related, strain 55989. As references for EAEC and EHEC, strains 042 and EDL933 were used, respectively. Interestingly, the *nanCMS*-operon which contains the genes for an outer membrane porin (NanC), a sialate mutarotase (NanM) and the Neu5,9Ac$_2$-esterase (NanS), was only present in the reference strains 042 and EDL933 but not in O104:H4 strains (Table 4). Moreover, NanS-p was present in multiple copies in O104:H4 strains, especially if these were *stx*2a-positive (Table 4). The NanS-p proteins can be divided in these strains into two groups (Fig. 3). The stx_{2a}-negative strains 55989 and 042 only express NanS-p proteins from group II, while the others have both, NanS-p from group I and II (Fig. 3). The *nanS*-p genes (*nanS*-p1, *nanS*-p2, *nanS*-p15, *nanS*-p23 and *nanS*-p29) which are downstream of *stx*2a or/and *stx*1a belong to group one and are closely related to each other (Fig. 3).

3.2. Expression of BTSA-His and NanS-p13-His

The *B. thetaiotaomicron* sialidase BTSA was characterized by Park et al. (2013) and was used in our study to cleave sialic acids from the sugar site chains of mucin from bovine submaxillary gland (see below). BTSA-His was recombinantly produced in a protein concentration of 0.2 mg/mL.

As a representative for the O104:H4 NanS-p proteins only *nanS*-p13 was cloned, and the expressed NanS-p13-His was further analyzed and used for control experiments (see below), since we recently demonstrated that EDL933 mutant strains could be complemented with different NanS-p proteins (Saile et al., 2016). NanS-p13-His was recombinantly produced in *E. coli* BL21(DE3)/pET-*nanS*-p13 and purified as described above and a protein concentration of 1.05 mg/mL was

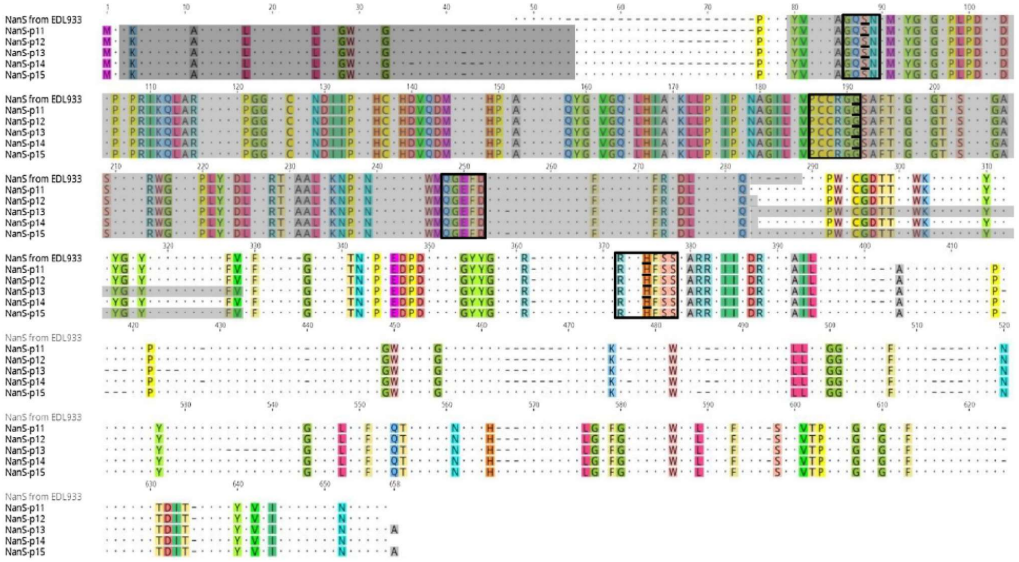

Fig. 2. ClustalW alignment of the LB226692 *nanS*-p alleles and *nanS* from EDL933. Dark grey and light grey boxes signals the DUF1737 and the SASA domain, respectively. The black-rimmed boxes indicate the motifs of the SGNH Type II hydrolase family.

achieved.

3.3. Detection of enzymatic activity of NanS-p13-His

An esterase activity of NanS-p13-His could be confirmed using 4-methylumbelliferyl-acetate as artificial test substrate. The highest esterase activity could be detected in Tris-HCl (pH 7.0) at 50 °C with a specific esterase activity of 35.8 ± 0.9 nkat/mg (Fig. 4).

3.4. Deletion mutagenesis

To investigate the functionality of the LB226692 NanS-p proteins, all *nanS*-p genes were deleted. Even after numerous attempts, deletion mutagenesis did not work in strain LB226692, putatively caused by a low transformation efficiency of only 10^4 to 10^5 cfu/µg pK18. Therefore, we used the O104:H4 *stx* $_{2a}$-negative derivative C227-11ϕcu. In contrast to strain LB226692, the transformation competence of C227-11ϕcu was approximately 10 times higher (10^5 –10^6 CFU per µg pK18). We confirmed the presence of *nanS*-p11, *nanS*-p12, *nanS*-p13 and *nanS*-p14 by PCR in this strain (data not shown). As expected, the genes *stx* $_{2a}$ and *nanS*-p15, corresponding to the *stx*2a-prophage were absent. The four remaining *nanS*-p genes, *nanS*-p11, *nanS*-p12, *nanS*-p13 and *nanS*-p14, were then successfully deleted in C227-11ϕcu as described above (data not shown), resulting in the mutant stains described in Table 1.

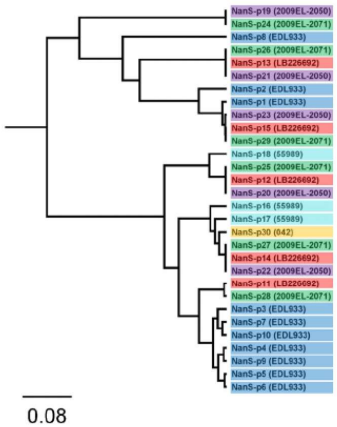

Fig. 3. Phylogenetic tree of NanS-p proteins from different *E. coli* genomes.

3.5. Growth on Neu5,9Ac $_2$ is dependent on the amount of nanS-p genes

Strain C227-11ϕcu, its mutant derivatives and the commensal strain AMC 198 were cultured in M9 minimal medium with Neu5,9Ac $_2$ as sole

Table 4
Presence (+) or absence (−) of the NanR regulated *nan*-operons and *stx*2a and the number of *nanS*-p in different *E. coli* strains.

Strain	Serotype	Pathotype	*nanATEK-yhjH*	*yjhBC*	*nanCMS*	No. of *nanS*-p	*stx* $_{2a}$
EDL933	O157:H7	EHEC	+	−	+	10	+
LB226692	O104:H4	EAHEC	+	−	−	5	+
55989	O104:H4	EAEC	+	−	−	3	−
2009EL-2050	O104:H4	EAHEC	+	−	−	5	+
2009EL-2071	O104:H4	EAHEC	+	−	−	6	+
042	O44:H18	EAEC	+	−	+	1	−

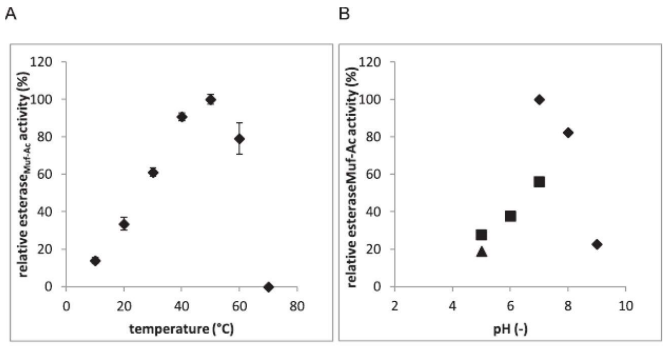

Fig. 4. Temperature (A) and pH (B) optima of recombinant NanS-p13-His. The temperature optimum was detected in Tris-HCl buffer (pH 7.0) and the pH optimum was measured at 40 °C in acetate (triangle), potassium phosphate (square) or in Tris-HCl (rhomb). A relative esterase$_{Muf-Ac}$ activity of 100% represents 35.8 ± 0.9 nkat/mg (A) or 36.2 ± 0.8 nkat/mg (B).

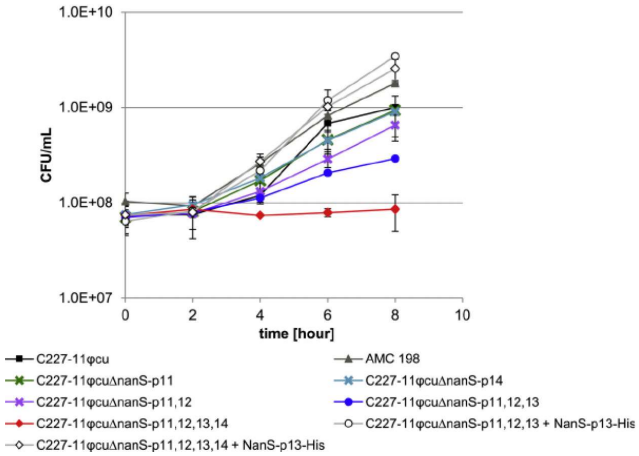

Fig. 5. Growth curves of commensal *E. coli* strain AMC 198, pathogenic *E. coli* strain O104:H4 C227-11Φcu and its *nanS-p* deletion mutants in M9 minimal medium with 0.4% (w/v) Neu5,9Ac$_2$. Experiments were performed in duplicate. Error bars depict the standard deviation of a biological duplicate with two technical replicates each.

carbon source and the respective growth curves were recorded. With step by step deletion of the different *nanS*-p alleles, an elongation of the generation time was detected ranging from 48 min for C227-11ϕcu to 84 min (C227-11ϕcuΔ*nanS*-p11), 92 min (C227-11ϕcuΔ*nanS*-p14), 106 min (C227-11ϕcuΔ*nanS*-p11,12), 137 min (C227-11ϕcuΔ*nanS*-p11,12,13) and up to 1272 min (C227-11ϕcuΔ*nanS*-p11,12,13,14) (Fig. 5). This result indicated that the number of *nanS*-p alleles of O104:H4 caused a gene-dose-effect on Neu5,9Ac$_2$. After complementation of C227-11ϕcuΔ*nanS*-p11,12,13 and C227-11ϕcu and C227-11ϕcu Δ*nanS*-p11,12,13,14 with NanS-p13-His, the generation times of 49 min and 63 min, respectively, were reduced to the wildtype level (Fig. 5).

Because of the presence of 4 *nanS*-p alleles in C227-11ϕcu, we expected better growth in M9 minimal medium with Neu5,9Ac$_2$ than for *E. coli* AMC 198 containing just a single *nanS* gene. Indeed, the generation time was 72 min, however, we could not measure significant differences in the cell counts until the time point of eight hours where actually AMC 198 revealed significant higher cell counts ($p < 0.05$) (Fig. 5).

In control experiments with all strains mentioned above, glucose or Neu5Ac was used as sole carbon source, and the optical density was determined at 600 nm over a 6 h period. No differences in growth behavior were detected, the tested strains grew equally well (data not shown).

3.6. E. coli O104:H4 strain C227-11 cu outgrows commensal E. coli on Neu5,9Ac$_2$

Following the hypothesis that Neu5,9Ac$_2$ represents a carbohydrate that offers pathogenic *E. coli* a competitive growth advantage towards commensal *E. coli,* we could demonstrate that in co-culturing experiments apathogenic strain AMC 198 was outgrown by C227-11ϕcu (Fig. 6A). The generation times for both strains increased as compared to the mono-culture experiment (see above). However, the difference was higher for AMC 198 with 40 min more than for C227-11ϕcu with an increase of the generation time by 14 min. The increase of generation times indicates a mutual influence of the strains in growth behavior, in which C227-11ϕcu is less influenced. This is indicated by the lower increase of its generation time, as compared to the increase in generation time of strain AMC198.

In contrast, if AMC 198 was co-cultured with C227-11ϕcuΔ*nanS*-p11,12,13 (Fig. 6B) or C227-11ϕcuΔ*nanS*-p11,12,13,14 (Fig. 6C), its generation times decreased by 12 or 17 min, respectively. This was surprising, because it indicated that AMC 198 benefited from the mutant strains. However, the generation times for C227-11ϕcuΔ*nanS*-p11,12,13 also decreased by 42 min and therefore AMC 198 was outgrown again. In contrast, AMC 198 could outgrow C227-11ϕcuΔ*nanS*-p11,12,13,14 in co-culture, although the mutant strain started to grow after 6 h probably using metabolic intermediates of AMC 198 (Fig. 6C). The superiority of C227-11Φcu towards AMC 198 could be explained by the quantity of *nanS*-p genes, a higher specificity of NanS-p toward

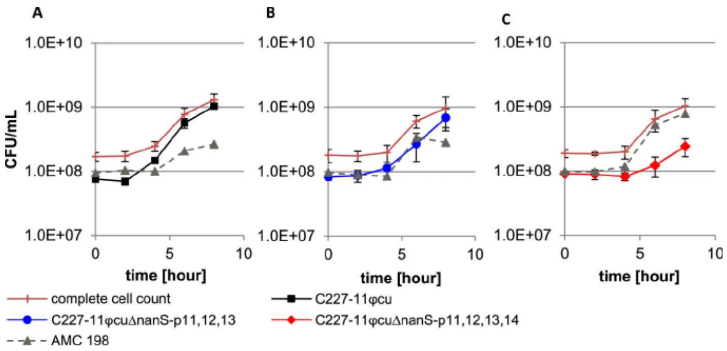

Fig. 6. Co-culture experiments of commensal *E. coli* strains AMC 198, C227-11Φcu (A) C227-11ΦcuΔ*nanS*-p11,12,13 (B), and C227-11ΦcuΔ*nanS*-p11,12,13,14 (C) in M9 minimal medium with 0.4% (w/v) Neu5,9Ac$_2$. Detection of CFU/mL was conducted on LB agar. The colony counts (CFU/mL) of C227-11Φcu (A) and mutant strains (B, C) were detected using LB-medium containing streptomycin. The colony counts (CFU/mL) for AMC 198 each were calculated by subtracting the mean of the total cell count with the mean of the cell counts of C227-11Φcu (A) or mutant strains (B, C). Error bars depict the standard deviation of a biological triplicate with two technical replicates each.

Neu5,9Ac$_2$ or by transcriptional incidents.

3.7. E. coli O104:H4 can not outgrow commensal E. coli using mucin as carbon source

Since Neu5,9Ac$_2$ occurs in high concentrations in mucin, we wanted to analyze whether the *nanS*-p genes also caused a growth advantage in the complex mucin as an energy source. In bovine submaxillary gland mucin, AMC 198 grew to higher cell counts in monoculture than O104:H4 (data not shown) and outgrew C227-11Φcu and C227-11ΦcuΔ*nanS*-p11,12,13,14 (Fig. 7A, B).

Sialic acids are normally connected to mucin with glycosidic bonds and are thought to be not available for *E. coli*, since they lack any genes for sialidase enzymes. We simulated the presence of a host microbiota-derived sialidase in the bacterial surrounding by adding BTSA-His to the mucin-containing medium (Fig. 7C, D). In general, we could repeat the results seen without BTSA (Fig. 7A, B). Therefore, the effect seen with the mono-substrate Neu5,9Ac$_2$ in the co-culture experiments (see 3.6) could not be seen with mucin as substrate. This demonstrates the complexity of mucin as a substrate and multiple metabolic pathways present in *E. coli*.

4. Discussion

In this study, we have demonstrated that the *E. coli* O104:H4 outbreak strain LB226692 contains five prophage-associated *nanS*-p alleles. These alleles are similar to those of *E coli* O157:H7 strain EDL933, but show significant differences in sequence.

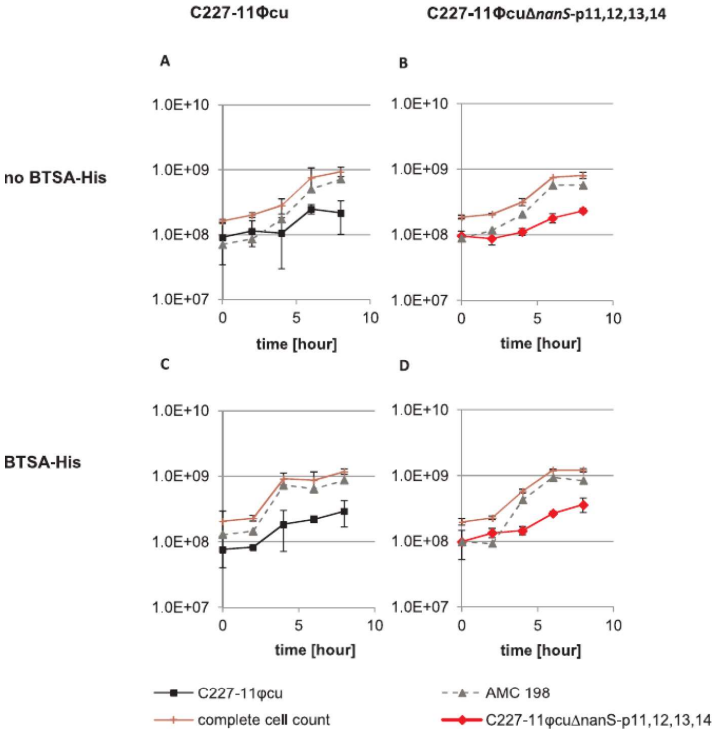

Fig. 7. Growth of commensal *E. coli* strain AMC 198, *E. coli* C227-11Φcu or C227-11ΦcuΔ*nanS*-p11,12,13,14 in co-culture in M9 minimal medium with 0.8% (w/v) mucin from bovine submaxillary gland without (A, B) or with (C, D) BTSA-His supplement in the medium. Co-culture of commensal *E. coli* strain AMC 198 and pathogenic *E. coli* O104:H4 C227-11Φcu (A, C) or C227-11ΦcuΔ*nanS*-p11,12,13,14 (B, D). Detection of total CFU per mL was performed using LB agar. Colony counts (CFU/mL) of C227-11Φcu (B) and C227-11ΦcuΔ*nanS*-p11,12,13,14 (C) were detected using LB with streptomycin. The colony counts (CFU/mL) for AMC 198 each were calculated by subtracting the mean of complete cell count with the mean of C227-11Φcu (B) or mutant strains (C).

However, proteins NanS-p15 and NanS-p1, which are encoded by the corresponding Stx2a-phages in LB226692 and EDL933 are 99% identical (Saile et al., 2016). The *nanS*-p genes are located in cryptic and functional prophages in 3′-direction of the antiterminator Q and upstream of the lysis genes. The only exception of this general observation is *nanS*-p12, which is located downstream of a holin- and an endolysin-encoding gene in the morphogenesis region of its corresponding prophage. That was surprising and we excluded putative sequencing artifacts by aligning the *nanS*-p12-associated prophage with the Blastn algorithm to already existing NCBI-sequences of O104:H4 isolates and found 100% consensus (data not shown). The Phaster analysis gave the *nanS*-p12-associated prophage the maximum score of 150 that mean that all genes of that region belong to an already known prophage (Zhou et al., 2011). From our sequence analysis, we rather think that this prophage is a composition of several phage elements and could be classified as cryptic prophage. The localization of the *nanS*-p genes might enable the genes to be co-expressed with the late phage genes and to be released from the bacterial cell upon phage-mediated cell lysis. However, there is evidence from the bioinformatical analysis (SecretomeP), that besides this, another secretion pathway may exist.

The different NanS-p proteins of LB226692 and other pathogenic *E. coli* could be assigned at least to two groups (Fig. 3). The enzymatic properties of NanS-p13 from LB226692 were a pH optimum of 7 and a temperature optimum of 50 °C (Fig. 4). That fit the optima we received for NanS-p1 and NanS-p2 of the same phylogenetic group, we analyzed in our previous study (Saile et al., 2016). In contrast, NanS-p4 of the other phylogenetic group had the same pH optimum but was nearly inactive at 50 °C. Its temperature optimum was 40 °C. Therefore, a phenotypical differences exist, that could offer the bacteria an adaptation to different environmental conditions. However, whether these genotypical differences has further implication on the function of NanS-p has to be analyzed in future.

Metabolism of sialic acids (sialometabolism) is very common in pathogenic and apathogenic *E. coli* and also in human pathogens such as *Vibrio cholerae, Streptococcus pneumoniae*, *Tannerella forsynthia* or *Cronobacter sakazakii*, which use them for competitive growth advantage (Almagro-Moreno and Boyd, 2009; Joseph et al., 2013; Kahya et al., 2017; Phansopa et al., 2015, Rangel et al., 2016). Therefore, it was surprising that O104:H4 serotypes lack the chromosomal *nanCMS*-operon, which basically is responsible for the catabolism of sialic acid derivatives. Similar to NanS, NanS-p proteins contain the SASA domain which is responsible for the esterase function.

The competition for nutrients of pathogens with metabolically related commensals had already been investigated in a mouse model with *Citrobacter rodentium*, a model organisms which cause A/E lesions in mice (Kamada et al., 2012). In our *in vitro* competition experiments the lack of *nanCMS* in O104:H4 had indeed no negative effects on the authority of C227-11ϕcu since C227-11ϕcu inhibited the growth of its "metabolic competitor" *E. coli* AMC 198, when grown on the sialic acid Neu5,9Ac$_2$. It is noteworthy, that a single *nanS*-p gene in *E. coli* O104:H4 was sufficient to outgrow strain AMC 198.

Sialic acids are highly abundant molecules at the contact area between bacteria and host (Robbe et al., 2004; Vimr et al., 2004). Cattle are the animals with the highest diversity of sialic acid derivatives (Schauer, 2004). Many bacteria present in the gut microbiota, such as *Bacteroides* spp., *Ruminoccocus gnavus* or *Bi dobacterium breve* are well adapted to mucin degradation by encoding sialidases (Crost et al., 2016; Egan et al., 2014; Juge et al., 2016; Robinson et al., 2017; Tailford et al., 2015). In this context, Robinson et al. (2017) recently demonstrated a synergistic effect between *Bacteroides* spp. and EHEC concerning sialic acid catabolism. We investigated the growth competition of O104:H4 and AMC 198 in bovine submaxillary gland mucin, where unexpectedly the commensal *E. coli* AMC 198 outgrew the O104:H4 wildtype with or without sialidase supplementation in the medium (Fig. 7). The reasons for this observation are currently not known. Cattle are presumably not the reservoir of *E. coli* O104:H4 and the sialic acid composition of bovine mucin differs from that of humans (Uber et al., 2006; Wieler et al., 2011). In our experiment, growth of both *E. coli* strains could be explained by the fact that mucin contain numerous carbon sources, and according to the metabolic flexibility of *E. coli*, thus metabolizing other mucin-derived carbohydrates or peptides at first and sialic acids at a later stage of growth. Another possible explanation is that the commercial bovine submaxillary gland mucin may contain free sugars and amino acids. In this case, *E. coli* would utilize free sugars and amino acids at first.

In conclusion, we showed that the catabolism of Neu5,9Ac$_2$ from *E. coli* O104:H4 differ from that of other *E. coli*, since the common *nanCMS*-operon, that initiate the catabolism of sialic acid derivatives, is absent. A high prevalence of the prophage-associated *nanS*-p genes, that are common in EHEC bacteria, was shown. We demonstrated for the first time, that the *nanCMS*-operon is not necessary for the catabolism of Neu5,9Ac$_2$, since O104:H4 could grew on that sialic derivative without the *nanCMS*-operon. Responsible for that phenomenon is the presence of multiple *nanS*-p genes in the genome. The key findings of our study were the importance of NanS-p proteins for O104:H4 in competition with the commensal *E. coli* strain AMC 198 for Neu5,9Ac$_2$.

Competing interest

The authors declare that they have no competing interests.

Funding

This work has been supported by grants of the Deutsche Forschungsgemeinschaft, grant Schm 1360/6-1.

Acknowledgments

We thank Prof. Alison D. O'Brien, Uniformed Services University of the Health Sciences, Bethesda, Maryland, United States for permission to use the *E. coli* O104:H4 *stx*$_{2a}$-negative derivative C227-11ϕcu, that we thankfully have received from Prof. Martina Bielaszewska, and Prof. Helge Karch, University of Münster, Germany. Moreover, we thank Heike Popovitsch, Claudia Lis and Markus Kranz for skilful technical assistance.

References

Almagro-Moreno, S., Boyd, E.F., 2009. Sialic acid catabolism confers a competitive advantage to pathogenic *Vibrio cholerae* in the mouse intestine. Infect. Immun. 77, 3807–3816.

Angiuoli, S.V., Matalka, M., Gussman, A., Galens, K., Vangala, M., Riley, D.R., Arze, C., White, J.R., White, O., Fricke, W.F., 2011. CloVR: a virtual machine for automated and portable sequence analysis from the desktop using cloud computing. BMC Bioinf. 12, 356. (Accessed 20 November 2017). http://clovr.org/.

Anon, 2017. Basic Local Alignment Search Tool. (Accessed 20 November 2017). https://blast.ncbi.nlm.nih.gov/Blast.cgi.

Anon, 2018a. SecretomeP 2.0. (Accessed 14 March 2018). http://www.cbs.dtu.dk/services/SecretomeP/.

Anon, 2018b. SignalP 4.1. (Accessed 14 March 2018). http://www.cbs.dtu.dk/services/SignalP/.

Arndt, D., Grant, J.R., Marcu, A., Sajed, T., Pon, A., Liang, Y., Wishart, D.S., 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 44, W16–21.

Bae, W.K., Lee, Y.K., Cho, M.S., Ma, S.K., Kim, S.W., Kim, N.H., Choi, K.C., 2006. A case of hemolytic uremic syndrome caused by *Escherichia coli* O104:H4. Yonsei Med. J. 47, 437–439.

Bendtsen, J.D., Kiemer, L., Fausbøll, A., Brunak, S., 2005. Non-classical protein secretion in bacteria. BMC Microbiol. 5, 58.

Bielaszewska, M., Mellmann, A., Zhang, W., Köck, R., Fruth, A., Bauwens, A., Peters, G., Karch, H., 2011. Characterisation of the *Escherichia coli* strain associated with an outbreak of heamolytic uraemic syndrome in Germany, 2011: a microbiological study. Lancet Infect. Dis. 11, 671–676.

Boisen, N., Ruiz-Perez, F., Scheutz, F., Krogfelt, K.A., Nataro, J.P., 2009. Short report: high prevalence of serine protease autotransporter cytotoxins among strains of enteroaggregative *Escherichia coli*. Am. J. Trop. Med. Hyg. 80, 294–301.

Brzuszkiewicz, E., Thürmer, A., Schuldes, J., Leimbach, A., Liesegang, H., Meyer, F.-D., Boelter, J., Petersen, H., Gottschalk, G., Daniel, R., 2011. 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.
Cherepanov, P.P., Wackernagel, W., 1995. 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.
Crost, E.H., Tailford, L.E., Monestier, M., Swarbreck, D., Henrissat, B., Crossman, L.C., Juge, N., 2016. The mucin-degradation strategy of *Ruminococcus gnavus*: the importance of intramolecular trans-sialidases. Gut microbes 7, 302–312.
Datsenko, K.A., Wanner, B.L., 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc. Natl. Acad. Sci. U. S. A. 97, 6640–6645.
Dautin, N., 2010. Serine protease autotransporters of enterobacteriaceae (SPATE): biogenesis and function. Toxins 2, 1179–1206.
Egan, M., O'Connell Motherway, M., Ventura, M., van Sinderen, D., 2014. Metabolism of sialic acid by *Bi dobacterium breve* UCC2003. Appl. Environ. Microbiol. 80, 4414–4426.
Finn, R.D., Clements, J., Arndt, W., Miller, B.L., Wheeler, T.J., Schreiber, F., Bateman, A., Eddy, S.R., 2015. HMMER web server: 2015 update. Nucleic Acids Res Web Server Issue 43:W30–W38.
Frank, C., Faber, M.S., Askar, M., Bernard Fruth, H., Gilsdorf, A., Hohle, A.M., Karch, H., Krause, G., Prager, R., Spode, A., Stark, K., Werber, D., 2011. Large and Ongoing Outbreak of Haemolytic Uraemic Syndrome, Germany, May 2011. Euro Surveill.
Harrington, S.M., Strauman, M.C., Abe, C.M., Nataro, J.P., 2005. Aggregative adherence fimbriae contribute to the inflammatory response of epithelial cells infected with enteroaggregative *Escherichia coli*. Cell. Microbiol. 7, 1565–1578.
Harrington, S.M., Sheikh, J., Henderson, I.R., Ruiz-Perez, F., Cohen, P.S., Nataro, J.P., 2009. The Pic protease of enteroaggregative *Escherichia coli* promotes intestinal colonization and growth in the presence of mucin. Infect. Immun. 77, 2465–2473.
Herold, S., Siebert, J., Huber, A., Schmidt, H., 2005. Global expression of prophage genes in *Escherichia coli* O157:H7 strain EDL933 in response to norfloxacin. Antimicrob. Agents Chemother. 49, 931–944.
Inoue, H., Nojima, H., Okayama, H., 1990. High efficiency transformation of *Escherichia coli* with plasmids. Gene 96, 23–28.
Joseph, S., Hariri, S., Masood, N., Forsythe, S., 2013. Sialic acid utilization by Cronobacter sakazakii. Microb. Inf. Exp. 3, 3.
Juge, N., Tailford, L., Owen, C.D., 2016. Sialidases from gut bacteria: a mini-review. Biochem. Soc. Trans. 44, 166–175.
Kahya, H.F., Andrew, P.W., Yesilkaya, H., 2017. Deacetylation of sialic acid by esterases potentiates pneumococcal neuraminidase activity for mucin utilization, colonization and virulence. PLoS Pathog. 13, e1006263.
Kamada, N., Kim, Y.-G., Sham, H.P., Vallance, B.A., Puente, J.L., Martens, E.C., Núñez, G., 2012. Regulated virulence controls the ability of a pathogen to compete with the gut microbiota. Science 336, 1325–1329.
Kaper, J.B., 1998. Enterohemorrhagic *Escherichia coli*. Curr. Opin. Microbiol. 1, 103–108.
King, L.A., Nogareda, F., Weill, F.-X., Mariani-Kurkdjian, P., Loukiadis, E., Gault, G., Jourdan-DaSilva, N., Bingen, E., Macé, M., Thevenot, D., Ong, N., Castor, C., Noël, H., van Cauteren, D., Charron, M., Vaillant, V., Aldabe, B., Goulet, V., Delmas, G., Couturier, E., Le Strat, Y., Combe, C., Delmas, Y., Terrier, F., Vendrely, B., Rolland, P., Valk de, H., 2012. Outbreak of Shiga toxin-producing *Escherichia coli* O104:H4 associated with organic fenugreek sprouts, France, June 2011. Clin. Infect. Dis. 54, 1588–1594.
Latif, H., Li, H.J., Charusanti, P., Palsson, B.Ø., Aziz, R.K., 2014. A gapless, unambiguous genome sequence of the enterohemorrhagic *Escherichia coli* O157:H7 strain EDL933. Genome Announc. 2, e00821–14.
Maniatis, T., Sambrook, J., Fritsch, E.F., 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring.
Mellmann, A., Bielaszewska, M., Köck, R., Friedrich, A.W., Fruth, A., Middendorf, B., Harmsen, D., Schmidt, M.A., Karch, H., 2008. Analysis of collection of hemolytic uremic syndrome-associated enterohemorrhagic *Escherichia coli*. Emerging Infect. Dis. 14, 1287–1290.
Mellmann, A., Harmsen, D., Cummings, C.A., Zentz, E.B., Leopold, S.R., Rico, A., Prior, K., Szczepanowski, R., Ji, Y., Zhang, W., McLaughlin, S.F., Henkhaus, J.K., Leopold, B., Bielaszewska, M., Prager, R., Brzoska, P.M., Moore, R.L., Guenther, S., Rothberg, J.M., Karch, H., 2011. Prospective genomic characterization of the German enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid next generation sequencing technology. PLoS One 6, e22751.
Mossoro, C., Glaziou, P., Yassibanda, S., Lan, N.T.P., Bekondi, C., Minssart, P., Bernier, C., Le Bouguénec, C., Germani, Y., 2002. Chronic diarrhea, hemorrhagic colitis, and hemolytic-uremic syndrome associated with HEp-2 adherent *Escherichia coli* in adults infected with human immunodeficiency virus in Bangui, Central African Republic. J. Clin. Microbiol. 40, 3086–3088.
Nübling, S., Eisele, T., Stöber, H., Funk, J., Polzin, S., Fischer, L., Schmidt, H., 2014. Bacteriophage 933W encodes a functional esterase downstream of the Shiga toxin 2a operon. Int. J. Med. Microbiol. 304, 269–274.
Ogura, H., Furuhata, K., Sato, S., Anazawa, K., Itoh, M., Shitori, Y., 1987. Synthesis of 9-O-acyl- and 4-O-acetyl-sialic acids. Carbohydr. Res. 167, 77–86.
Park, K.H., Kim, M.G., Ahn, H.J., Lee, D.H., Kim, J.H., Kim, Y.W., Woo, E.J., 2013. Structural and biochemical characterization of the broad substrate specificity of Bacteroides thetaiotaomicron commensal sialidase. Biochim. Biophys. Acta 1834, 1510–1519.
Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8, 785–786.
Phansopa, C., Kozak, R.P., Liew, L.P., Frey, A.M., Farmilo, T., Parker, J.L., Kelly, D.J., Emery, R.J., Thomson, R.I., Royle, L., Gardner, R.A., Spencer, D.I.R., Stafford, G.P., 2015. Characterization of a sialate-O-acetylesterase (NanS) from the oral pathogen *Tannerella forsythia* that enhances sialic acid release by NanH, its cognate sialidase. Biochem. J. 472, 157–167.
Polzin, S., Huber, C., Eylert, E., Elsenhans, I., Eisenreich, W., Schmidt, H., 2013. Growth media simulating ileal and colonic environments affect the intracellular proteome and carbon fluxes of enterohemorrhagic *Escherichia coli* O157:H7 strain EDL933. Appl. Environ. Microbiol. 79, 3703–3715.
Pridmore, R.D., 1987. New and versatile cloning vectors with kanamycin-resistance marker. Gene 56, 309–312.
Rangarajan, E.S., Ruane, K.M., Proteau, A., Schrag, J.D., Valladares, R., Gonzalez, C.F., Gilbert, M., Yakunin, A.F., Cygler, M., 2011. Structural and enzymatic characterization of NanS (YjhS): a 9-O-acetyl *N*-acetylneuraminic acid esterase from *Escherichia coli* O157:H7. Protein Sci. 20, 1208–1219.
Rangel, A., Steenbergen, S.M., Vimr, E.R., 2016. Unexpected diversity of *Escherichia coli* sialate O-acetyl esterase NanS. J. Bacteriol. 198, 2803–2809.
Rasko, D.A., Webster, D.R., Sahl, J.W., Bashir, A., Boisen, N., Scheutz, F., Paxinos, E.E., Sebra, R., Chin, C.-S., Iliopoulos, D., Klammer, A., Peluso, P., Lee, L., Kislyuk, A.O., Bullard, J., Kasarskis, A., Wang, S., Eid, J., Rank, D., Redman, J.C., Steyert, S.R., Frimodt-Møller, J., Struve, C., Petersen, A.M., Krogfelt, K.A., Nataro, J.P., Schadt, E.E., Waldor, M.K., 2011. Origins of the *E. coli* strain causing an outbreak of hemolytic-uremic syndrome in Germany. N. Engl. J. Med. 365, 709–717.
Rauw de, K., Vincken, S., Garabedian, L., Levtchenko, E., Hubloue, I., Verhaegen, J., Craeghs, J., Glupczynski, Y., Mossong, J., Piérard, D., 2014. Enteroaggregative shiga toxin-producing *Escherichia coli* of serotype O104:H4 in Belgium and Luxembourg. New Microbes New Infect. 2, 138–143.
Robbe, C., Capon, C., Coddeville, B., Michalski, J.-C., 2004. Structural diversity and specific distribution of *O*-glycans in normal human mucins along the intestinal tract. Biochem. J. 384, 307–316.
Robert Koch-Institut, 2012. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2011. Robert Koch-Institut, Berlin (Accessed 20 November 2017). https://www.rki.de/DE/Content/Infekt/Jahrbuch/Jahrbuch_2011. pdf?_blob=publicationFile.
Robinson, L.S., Lewis, W.G., Lewis, A.L., 2017. The sialate *O*-acetylesterase EstA from gut *Bacteroidetes* species enables sialidase-mediated cross-species foraging of 9-O-acetylated sialoglycans. J. Biol. Chem. 292, 11861–11872.
Saile, N., Voigt, A., Kessler, S., Stressler, T., Klumpp, J., Fischer, L., Schmidt, H., 2016. *Escherichia coli* O157:H7 strain EDL933 harbors multiple functional prophage-associated genes necessary for the utilization of 5-N-acetyl-9-O-acetyl neuraminic acid as a growth substrate. Appl. Environ. Microbiol. 82, 5940–5950.
Schauer, R., 2004. Sialic acids: fascinating sugars in higher animals and man. Zoology 107, 49–64.
Scheutz, F., Nielsen, E.M., Frimodt-Møller, J., Boisen, N., Morabito, S., Tozzoli, R., Nataro, J.P., Caprioli, A., 2011. Characteristics of the Enteroaggregative Shiga Toxin/verotoxin-producing *Escherichia Coli* O104:H4 Strain Causing the Outbreak of Haemolytic Uraemic Syndrome in Germany, May to June 2011. Euro Surveill., pp. 2011 (16, pii: 19889).
Steenbergen, S.M., Jirik, J.L., Vimr, E.R., 2009. YjhS (NanS) is required for *Escherichia coli* to grow on 9-O-acetylated N-acetylneuraminic acid. J. Bacteriol. 191, 7134–7139.
Studier, F.W., Moffatt, B.A., 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. 189, 113–130.
Tailford, L.E., Owen, C.D., Walshaw, J., Crost, E.H., Hardy-Goddard, J., Le Gall, G., Vos, W.M., de Taylor, G.L., Juge, N., 2015. Discovery of intramolecular trans-sialidases in human gut microbiota suggests novel mechanisms of mucosal adaptation. Nat. Commun. 6, 7624.
Tietze, E., Dabrowski, P.W., Prager, R., Radonic, A., Fruth, A., Auraß, P., Nitsche, A., Mielke, M., Flieger, A., 2015. Comparative genomic analysis of two novel sporadic Shiga toxin-producing *Escherichia coli* O104:H4 strains isolated 2011 in Germany. PLoS One 10, e0122074.
Uber, A.P., Trabulsi, L.R., Irino, K., Beutin, L., Ghilardi, A.C.R., Gomes, T.A.T., Liberatore, A.M.A., Castro, A.F.P., de Elias, W.P., 2006. Enteroaggregative *Escherichia coli* from humans and animals differ in major phenotypical traits and virulence genes. FEMS Microbiol. Lett. 256, 251–257.
Vimr, E.R., Kalivoda, K.A., Deszo, E.L., Steenbergen, S.M., 2004. Diversity of microbial sialic acid metabolism. Microbiol. Mol. Biol. Rev. 68, 132–153.
Vimr, E.R., 2013. Unified theory of bacterial sialometabolism: how and why bacteria metabolize host sialic acids. ISRN Microbiol. 2013, 816713.
Wieler, L.H., Semmler, T., Eichhorn, I., Antao, E.M., Kinnemann, B., Geue, L., Karch, H., Guenther, S., Bethe, A., 2011. No evidence of the Shiga toxin-producing E. coli O104:H4 outbreak strain or enteroaggregative *E. coli* (EAEC) found in cattle faeces in northern Germany, the hotspot of the 2011 HUS outbreak area. Gut Pathog. 3, 17.
Xu, J., Bjursell, M.K., Himrod, J., Deng, S., Carmichael, L.K., Chiang, H.C., Hooper, L.V., Gordon, J.I., 2003. A genomic view of the human-Bacteroides thetaiotaomicron symbiosis. Science 299, 2074–2076.
Zangari, T., Melton-Celsa, A.R., Panda, A., Boisen, N., Smith, M.A., Tatarov, I., Tolla, L.J., de Nataro, J.P., O'Brien, A.D., 2013. Virulence of the Shiga toxin type 2-expressing *Escherichia coli* O104:H4 German outbreak isolate in two animal models. Infect. Immun. 81, 1562–1574.
Zhou, Y., Liang, Y., Lynch, K.H., Dennis, J.J., Wishart, D.S., 2011. PHAST: a fast phage search tool. Nucleic Acids Res. 39, W347–W352.

Kapitel 4

De-*O*-Acetylation of mucin-derived sialic acids by recombinant NanS-p esterases of *Escherichia coli* O157:H7 strain EDL933

Stefanie Feuerbaum[1], Nadja Saile[1], Gottfried Pohlentz[2], Johannes Müthing[2], Herbert Schmidt[1]

[1] Department of Food Microbiology and Hygiene, Institute of Food Science and Biotechnology, University of Hohenheim, Stuttgart, Germany

[2] Institute for Hygiene, University of Münster, Münster, Germany

Publiziert in International Journal of Medical Microbiology, Band 308, Seiten 1113-1120.

DOI: 10.1016/j.ijmm.2018.10.001. Mit freundlicher Genehmigung von Elsevier.

International Journal of Medical Microbiology 308 (2018) 1113–1120

Contents lists available at ScienceDirect

International Journal of Medical Microbiology

journal homepage: www.elsevier.com/locate/ijmm

De-*O*-Acetylation of mucin-derived sialic acids by recombinant NanS-p esterases of *Escherichia coli* O157:H7 strain EDL933

S. Feuerbaum[a,1], N. Saile[a,1], G. Pohlentz[b], J. Müthing[b], H. Schmidt[a,*]

[a] Department of Food Microbiology and Hygiene, Institute of Food Science and Biotechnology, University of Hohenheim, Stuttgart, Germany
[b] Institute for Hygiene, University of Münster, Germany

ARTICLE INFO

Keywords:
O-acety-l-esterase
Escherichia coli O157:H7
Bovine submaxillary gland mucin
De-*O*-acetylation
Neu5A
Neu5Ac2
NanS-p
NanS

ABSTRACT

Enterohemorrhagic *Escherichia coli* (EHEC) O157:H7 strain EDL933 encodes the single chromosomal 9-*O*-acetylesterase NanS, and several copies of prophage-encoded 9-*O*-acetylesterases (NanS-p). These enzymes have recently been shown to cleave 5-*N*-acetyl-9-*O*-acetyl neuraminic acid (Neu5,9Ac$_2$) to yield de-*O*-acetylated Neu5Ac, the latter of which may serve as a carbon and/or nitrogen source.

In the current study, we investigated the NanS- and NanS-p-mediated digestion of synthetic *O*-acetylated neuraminic acids and bovine submaxillary glands mucin (BSM)-derived *O*-acetylneuraminic acids by high-performance thin-layer chromatography (HPTLC) and nano electrospray ionization mass spectrometry (nanoESI MS). Initial HPTLC analyses showed the expected activity of NanS and NanS-p variants for Neu5,9Ac$_2$. However, all tested enzymes were unable to de-*O*-acetylate 5-*N*-acetyl-4-*O*-acetylneuraminic acid (Neu5,4 Ac$_2$) in our test system. The nanoESI MS analysis of neuraminic acids after treatment of BSM with NanS-p gave evidence that NanS-p variants of EHEC O157:H7 strain EDL933 cleave off *O*-acetyl groups from mono-, di-, and tri-*O*-acetylated Neu5Ac and *N*-glycolylneuraminic acid (Neu5Gc), regardless of the carbon positions C7, C8 or C9 of the acetate esters. This enzyme activity leads to neuraminidase-accessible Neu5Ac and Neu5Gc on mucin glycans.

Moreover, we could demonstrate by HPTLC analyses that recombinant *Bacteroides thetaiotaomicron* sialidase (BTSA-His) was able to cleave Neu5Ac and Neu5,9Ac$_2$ from BSM and that the combination of BTSA-His with both NanS-His and NanS-p-His derivatives enhanced the release of de-*O*-acetylated core Neu5Ac and Neu5Gc from mammalian mucin *O*-glycans.

Growth experiments with EHEC wildtype strain EDL933, its *nanS* and *nanS/nanS*-p1a-p7 mutant and exogenous BTSA-His in BSM demonstrated that the presence of BTSA-His enhanced growth of EDL933 and the *nanS* deletion mutant but not the *nanS/nanS*-p1a-p7 mutant.

Thus, we hypothesize that the expression of sialic acid *O*-acetylesterases with a broad specificity could be an advantage in competition with the gut microbiota for nutrients and facilitate EHEC colonization in the human large intestine.

1. Introduction

Infections with enterohemorrhagic *Escherichia coli* (EHEC) can cause hemorrhagic colitis (HC) and the life-threatening hemolytic-uremic syndrome (HUS) (Nataro and Kaper, 1998). One of the natural reservoirs of EHEC strains is the intestinal tract of ruminants, especially cattle, where they can survive in the presence of mucus-derived carbohydrates (Aperce et al., 2014; Bertin et al., 2013; Fox et al., 2009). EHEC O157:H7 strains have caused numerous cases and outbreaks during the last decades (Page and Liles, 2013) Most strains of this serotype express Shiga toxin (Stx) 1 and/or Stx2 with their different variants and the locus of enterocyte effacement (LEE), encoding a type III secretion system (Scheutz et al., 2011; Stevens and Frankel, 2014). The first reported EHEC O157:H7 outbreak occurred in 1982 where at least 47 people developed diarrhea in Michigan and Oregon after consuming contaminated meat in a fast-food restaurant (Riley et al., 1983). One of the strains that have been isolated during this outbreak was EDL933 (O'Brien et al., 1993). Up to now, this is one of the best-characterized EHEC O157:H7 strains worldwide and was therefore used in the current study.

* Corresponding author at: Department of Food Microbiology and Hygiene, Institute of Food Science and Biotechnology, University of Hohenheim, Garbenstrasse 28, 70599, Stuttgart, Germany.
E-mail address: herbert.schmidt@uni-hohenheim.de (H. Schmidt).
[1] Stefanie Feuerbaum and Nadja Saile contributed equally to this work.

https://doi.org/10.1016/j.ijmm.2018.10.001
Received 6 July 2018; Received in revised form 27 September 2018; Accepted 5 October 2018

The primary site of infection in the human body is the large intestine, where EHEC cells adhere to the intestinal epithelium, damage this cellular barrier by the action of type III effector proteins and translocate Stx into the blood system (Nguyen and Sperandio, 2012). Infecting EHEC strains may evade host defence, gain access to growth substrates and overcome colonization resistance mechanisms (Stecher and Hardt, 2011). They replicate in the colon during the course of infection, but up to now, the mechanisms of competition with the microbiota and carbon nutrition are only poorly understood. An important carbon source for EHEC are breakdown products of the gastrointestinal mucus that are produced by the host in high amounts each day (Conway et al., 2004). Many commensal gut microorganisms attain required *C*- and *N*-sources from mucin glycoproteins (Johansson et al., 2009, 2011; Marcobal et al., 2013). Endowed with various glycoside hydrolases and exploiting synergistic mechanisms between them they are well adapted to grow in the mucus environment (Marcobal et al., 2013).

A number of EHEC O157:H7 virulence factors seems to play primary roles in penetration of mucus or adhesion to the mucus-coated colonic epithelium such as the mucinase StcE (Hews et al., 2017) or the H7 flagella (Erdem et al., 2007). To get access to the limited nutrients in large intestine, replication of EHEC strains depends on mechanisms to compete at least for one of the mucus-derived carbohydrates or to be otherwise metabolically flexible (Conway and Cohen, 2015). Commensal non-pathogenic *E. coli* represent presumably the major competitors in the gut, because they exhibit a similar genetic background as EHEC do. Thus, the expression of additional genes involved in sugar catabolism, which are located on genetic mobile elements in EHEC, is imperative for the assertiveness of EHEC towards commensal *E. coli*. Mucus-derived sugars as fucose, galactose, *N*-acetylglucosamine, *N*-acetylgalactosamine or *N*-acetylneuraminic acid (Neu5Ac) can be utilized by EHEC as well as by commensal *E. coli* (Aperce et al., 2014; Chang et al., 2004; Fabich et al., 2008).

Neu5Ac is an α-keto sugar with a nine-carbon backbone, typically found terminally linked to glycan chains of mammalian mucins. Neu5Ac is highly abundant in glycans of the large intestine (Robbe et al., 2003), where *E. coli* preferentially colonizes, compared to the small intestine (Barnett Foster, 2013; Torres et al., 2005). *N*-glycolylneuraminic acid (Neu5Gc) is the hydroxylated derivative of Neu5Ac carrying an OH-group in the *N*-acetyl chain at carbon C5. Neu5Gc is generated by the enzyme CMP-Neu5Ac-hydroxylase, which is expressed as functional enzyme in all mammals except humans, who produce an inactive enzyme due a deletion in the corresponding chromosome (Peri et al., 2018). Although humans lack a functioning CMP-Neu5Ac-hydroxylase, Neu5Gc has been found in low quantities in human mucins most likely due to catabolic incorporation from nutritional Neu5Gc derived from red meat (Varki and Schauer, 2009). Furthermore, neuraminic acids exhibit a large heterogeneity in the gut with respect to structural modifications of their hydroxyl groups. A very common modification is the *O*-acetylation with up to four *O*-acetyl residues at carbon atoms C4, C7, C8 and C9 (Varki and Schauer, 2009; Vimr, 2013). Neu5Ac is the major sialic acid present in mucin, followed by the mono-*O*-acetylated Neu5Ac-variants Neu5,9Ac_2, Neu5,7Ac_2 and Neu5,8Ac_2. Less common are the di-*O*-acetylated Neu5Ac-species Neu5,8,9Ac_3 and Neu5,7,9Ac_3 (Robbe et al., 2003).

Nonpathogenic and pathogenic *E. coli* strains can encode the three chromosomal operons *nanATEK-yhcH*, *nanCMS* (formerly *yjhATS*), and *yjhBC*, which are involved in sialic acid metabolism. Neu5Ac inactivates the repressor NanR, whereby the *nan*-operons are activated and sialic acids are catabolized (Rangarajan et al., 2011). While Neu5Ac is directly taken up by the transporter NanT, *O*-acetylated sialic acids have to be deacetylated first (Steenbergen et al., 2009). NanS is an *N*-acetyl-9-*O*-acetylneuraminic acid esterase, which deacetylates mono-*O*-acetylated Neu5,9Ac_2 to Neu5Ac, which is then transported by NanT and utilized as an energy source. It has been shown that NanS is also able to deacetylate mono-*O*-acetylated Neu5,8Ac_2, but not the Neu5,7Ac_2 or Neu5,4 Ac_2 isomeric structures (Steenbergen et al., 2009).

Expressing NanS, commensal non-pathogenic *E. coli* are able to utilize *O*-acetylated sialic acids as carbon sources. NanS belongs to the large family of SGNH serine proteases, which are not only found in commensal bacteria, but also in a number of pathogens that produce sialate *O*-acetylesterases (Corfield et al., 1992, 1993; Phansopa et al., 2015). Vimr (2013) have extensively reviewed the bacterial sialometabolism and reported that many enteropathogenic *E. coli* strains harbor further *nanS*-homologs.

The pathogenic *E. coli* O157:H7 strain EDL933 encodes several *nanS*-homologous alleles (*nanS*-p), which are localized in prophage genomes in the region of late gene expression (Saile et al., 2016). Recently, the enzymatic characteristics of recombinantly expressed NanS-p1, NanS-p2, and NanS-p4 of strain EDL933 were determined and an *N*-acetyl-9-*O*-acetylneuraminic acid esterase activity was detected for the three enzymes. The three variants were investigated as representatives of the EDL933 NanS-p and were selected due to phylogenetic reasons (Saile et al., 2016). For analysis of the role of NanS-p, several *nanS* and *nanS*-p deletion mutants have been constructed up to a supermutant that did not contain any *nanS* or *nanS*-p genes, and their role in substrate utilization has been demonstrated (Saile et al., 2016). Compared to *nanS*, the *nanS*-p genes harbour additional nucleotides downstream of the SGNH hydrolase-like domain and a DUF1737 domain. The function of the DUF1737 domain is still unknown (Rangel et al., 2016; Saile et al., 2016).

In the human large intestine, commensal mucin-degrading species of the phyla *Bacteroidetes*, *Firmicutes*, *Actinobacteria* and *Verrucomicrobia* can cleave bound sialic acids from glycans by intrinsic sialidases, providing free nutritional sialic acids for sialidase-negative *E. coli* (Robinson et al., 2017; Tailford et al., 2015; Vimr et al., 2004). Sialidases are substrate-specific and the decoration of sialic acids with one or more *O*-acetyl residues prevents sialidases from enzymatic cleavage. The activity of sialidases increases upon removal of *O*-acetyl groups from sialic acids by sialate *O*-acetylesterases (Phansopa et al., 2015). *B. thetaiotaomicron*, for example, represents a commensal of the gut, that produces a number of glycoside hydrolases such as the *B. thetaiotaomicron* sialidase BTSA (Park et al., 2013), enabling this microorganism to release terminally bound sialic acids from mucin.

The aim of the current study was to further characterize the enzymatic properties of recombinant NanS and NanS-p derivatives to cleave *O*-acetyl residues from *O*-acetylated neuraminic acids and BSM, also in combination with recombinant BTSA. Moreover, growth experiments with EHEC wildtype strains and *nanS*-p mutants were performed to study the influence of the enzymes *in vitro*.

2. Materials and methods

2.1. Bacterial strains

Bacterial strains used in this study are listed in Table 1.

2.2. Cloning and expression techniques

E. coli O157:H7 strain EDL933 was grown on LB agar (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, pH 7) at 37 °C with shaking at 180 rpm. Amplification of *nanS* from EDL933 was performed by designing primers Z5905_NdeI_for (5'-ccc*CATATG*AACGCAATAAT ATCGCCTG-3') and Z5905_XhoI_rev (5'-aaa*CTCGAG*CCTTTCGCGCCA AAACTGC-3') using the NCBI reference sequence NZ_CP015855 (Saile et al., 2016). PCR conditions were as follows: after initial DNA denaturation for 5 min at 95 °C, 30 cycles of each 30 s 94 °C, 30 s 66 °C, and 120 s 72 °C were carried out. After a final extension step of 5 min at 72 °C, the PCR products were cooled down to 4 °C and further processed. The PCR product was then digested with the restriction enzymes *Nde*I and *Xho*I (Thermo Scientific, USA), ligated into expression vector pET-22b(+), transformed into laboratory *E. coli* strain BL21(DE3), and correct insertion of the PCR product was confirmed by DNA sequencing

Table 1
Characteristics of bacterial strains used in this study.

Bacterial strains	Relevant characteristics	Reference or source
E. coli O157:H7 EDL933	*nanS*$^{+}$, *nanS*-p1a^{+} to *nanS*-p7^{+}	(O'Brien et al., 1993)
E. coli O157:H7 EDL933Δ*nanS*	Deletion of *nanS*	(Saile et al., 2016)
E. coli O157:H7 EDL933Δ*nanS*Δ*nanS*-p1a-p7	Deletion of *nanS*, *nanS*-p1a, *nanS*-p2, *nanS*-p3, *nanS*-p4, *nanS*-p5, *nanS*-p6, *nanS*-p7	(Saile et al., 2016)
E. coli BL21(DE3)	F– *ompT hsdSB*(rB^{-},mB^{-}) *gal dcm* (DE3)	(Studier and Moffatt, 1986)
E. coli BL21(DE3)/pET-22b(+)	contains plasmid pET-22b(+) (Novagene, Merck)	(Nübling et al., 2014)
E. coli BL21(DE3)/pET-Z1466-his	Recombinant 933 Wp42 (syn. NanS-p1)	(Nübling et al., 2014)
E. coli BL21(DE3)/pET-*nanS*-p2-his	Recombinant NanS-p2	(Saile et al., 2016)
E. coli BL21(DE3)/pET-*nanS*-p4-his	Recombinant NanS-p4	(Saile et al., 2016)
E. coli BL21(DE3)/pET-BT_0455-his	Recombinant BTSA (without signal sequence)	(Saile et al., 2018)
E. coli BL21(DE3)/pET-*nanS*-his	Recombinant NanS	This study

as described previously (Nübling et al., 2014).

Expression, verification of protein purity and determination of the concentrations of BTSA-His, NanS-His, NanS-p1-His, NanS-p2-His and NanS-p4-His by SDS-PAGE and Bradford staining were performed as recently described (Nübling et al., 2014; Saile et al., 2016, 2018).

2.3. Growth of EDL933 and its O-acetyl esterase knockout derivatives in bovine submaxillary gland mucin supplemented with BTSA-His

E. coli O157:H7 strain EDL933, EDL933Δ*nanS* and EDL933Δ*nanS*Δ*nanS*-p1a-p7 were incubated for 24 h at 37 °C and 180 rpm in M9 minimal medium (Sambrook and Russell, 2001) containing 0.8% (w/v) mucin from bovine submaxillary gland (BSM, cat. 84195-52-8, Merck Millipore, Germany). The initial optical density (OD_{600}) was 0.1 for each culture. Liquid culture of EDL933Δ*nanS*Δ*nanS*-p1a-p7, supplemented with 4 µg/ml NanS-p4-His, was used as a control.

To compare the effect of recombinant BTSA-His on growth of EHEC strains, the cultures were grown in parallel, whereby one of the parallel culture each contained 3.86 µg/ml BTSA-His. All experiments were performed in triplicate with two technical replicates, respectively. Statistical significance was determined using Student's *t* test.

2.4. Enzymatic digestion of BSM

One mg/ml of BSM dissolved in 50 mM Tris-HCl pH 7.5 was incubated with 5 µg NanS-His, NanS-p1-His, NanS-p2-His or NanS-p4-His for 2 h at 25 °C. Then another 5 µg of enzyme was added (10 µg of NanS-p equates to 0.15 µM; 10 µg of NanS equates to 0.27 µM) and incubation was prolonged for 2 h. In addition, 1 mg/ml BSM was incubated with 0.96 µg/mg (0.02 µM) BTSA-His for 4 h at 25 °C. Digestions of BSM with NanS-His/NanS-p-His and BTSA-His were carried out similar to the BSM digestion with NanS-His/NanS-p-His but contained in addition 0.96 µg/mg BTSA-His. BSM was incubated for 4 h at 25 °C without enzyme and used as negative control. All reactions were stopped after the incubation time by heat inactivation at 95 °C for 10 min.

2.5. Enzymatic de-O-acetylation of Neu5,9Ac$_2$ and Neu5,4 Ac$_2$

One hundred µg of mono-*O*-acetylated Neu5,9Ac$_2$ (provided by Wolfgang Fessner and Ning He, Technical University Darmstadt, Darmstadt, Germany) or mono-*O*-acetylated Neu5,4 Ac$_2$ (Applied Biotech GmbH, Austria, no. S-4-S-10) were dissolved in 100 µl of 50 mM Tris-HCl pH 7.5 and incubated each with 5 µg NanS-His, NanS-p1-His, NanS-p2-His or NanS-p4-His at 25 °C for 1 h. Another 5 µg of enzyme (in total 1.5 µM of NanS-p and 2.7 µM of NanS) were added to each reaction and incubation was prolonged for 1 h. After incubation the enzymes were heat inactivated at 95 °C for 10 min. As a negative control 100 µg of sialic acids (Neu5Ac, Neu5,9Ac$_2$, Neu5,4 Ac$_2$) were incubated for 2 h at 25 °C without enzymes.

2.6. High-performance thin-layer chromatography (HPTLC) analysis of sialic acids

HPTLC analysis was performed with glass-backed HPTLC plates (20 cm x 10 cm) coated with silica gel 60 F_{254} (Merck, Germany). Three to 6 µl of heat-inactivated enzymatic digestions and Neu5Ac, Neu5,9Ac$_2$ and Neu5,4 Ac$_2$ standards were applied onto plates as a fine 8 mm wide streak using the Automatic TLC Sampler 4 (ATS4, CAMAG, Switzerland) equipped with a 100 µl Hamilton® syringe. After sample application the plates were developed in a twin trough developing chamber (CAMAG) filled with n-butanol:H_2O:acetic acid (5:2.5:2.5, each by vol.) as a mobile phase. Plates were first air-dried and then dipped in diphenylamine-aniline-phosphoric acid (DAPA, Carl Roth, Germany) reagent for one second and kept on a TLC plate heater (CAMAG, Switzerland) at 110 °C until colored bands became apparent. Visualization was done under UV light (366 nm) using a TLC visualizer (CAMAG). The chromatograms were processed with winCATS software (CAMAG).

2.7. Sample preparation for nanoESI MS

BSM samples were incubated with NanS-His, NanS-p1-His, NanS-p2-His, NanS-p4-His, or without enzyme followed by treatment with 12.5% (v/v) acetic acid for 2 h at 95 °C. After chilling to ambient temperature the samples were dried *in vacuo* and dissolved in 40% methanol containing 0.5% formic acid for MS analysis. Samples treated with BTSA-His (*vide supra*) were spiked with an equal volume of methanol an incubated on ice for 30 min. After centrifugation at 15,800 x *g* the supernatants were analyzed by nanoESI MS.

2.8. nanoESI mass spectrometry analysis

Mass spectrometry (MS) analyses were performed using a SYNAPT G2-S mass spectrometer (Waters, Manchester, UK) equipped with a Z-spray source in the positive ion sensitivity mode. Typical source settings were: temperature 80 °C, capillary voltage 0.8 kV, sampling cone voltage 20 V, and offset voltage 50 V (Steil et al., 2018).

3. Results

3.1. Recombinant EHEC O-acetylesterases NanS, NanS-p1, NanS-p2 and NanS-p4 deacetylate mono-O-acetylated Neu5,9Ac$_2$, but not Neu5,4 Ac$_2$

The native *O*-acetylesterases NanS and NanS-p have been shown to remove the 9-*O*-acetyl group from Neu5,9Ac$_2$ (Nübling et al., 2014; Saile et al., 2016). In order to explore the activity and specificity of purified NanS and NanS-p, we cloned His-tagged *nanS*, *nanS*-p1, *nanS*-p2, and *nanS*-p4 and expressed the enzymes in laboratory *E. coli* strain BL21 (DE3) following IPTG induction as described above. Cells were lysed and recombinant enzymes were purified with Ni-NTA-columns,

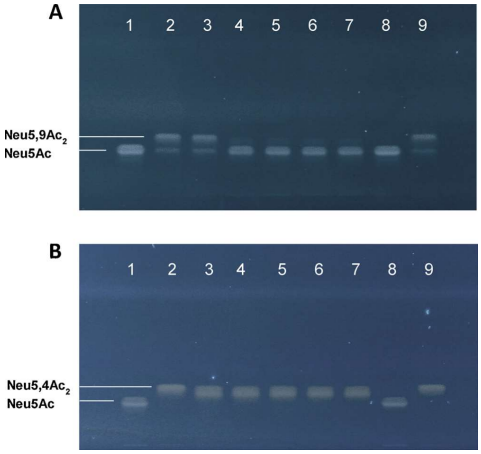

Fig. 1. (A) HPTLC analysis of mono-*O*-acetylated Neu5,9Ac_2 and (B) Neu5,4 Ac_2 incubated with NanS-p EHEC *O*-acetylesterases. Three µl of 1 mg/mL Neu5Ac (Fig. 1A,B, lanes 1,8), Neu5,9Ac_2 (Fig. 1A, lanes 2,9) and Neu5,4 Ac_2 (Fig. 1B, lanes 2,9) preparations each, were applied onto HPTLC plates as controls. In addition, three µl of 1 mg/ml Neu5,9Ac_2 (Fig. 1A, lane 3), and Neu5,4 Ac_2 (Fig. 1B, lane 3) containing reaction buffer without enzymes were applied as buffer controls. Neu5,9Ac_2 was digested with NanS-His (Fig.1A, lane 4), NanS-p1-His (Fig.1A, lane 5), NanS-p2-His (Fig.1A, lane 6) and NanS-p4-His (Fig.1A, lane 7). Neu5,4 Ac_2 was digested with NanS-His (Fig.1B, lane 4), NanS-p1-His (Fig. 1B, lane 5), NanS-p2-His (Fig. 1B, lane 6), and NanS-p4-His (Fig. 1B, lane 7).

followed by determining their purity and concentration by SDS-PAGE and Bradford staining. NanS-His, NanS-p1-His, NanS-p2-His and NanS-p4-His were incubated for 2 h with mono-*O*-acetylated Neu5,9Ac_2 or Neu5,4 Ac_2; samples of Neu5,9Ac_2 and Neu5,4 Ac_2 without enzymes served as controls. After enzyme treatment, we qualitatively analysed the extent of *O*-acetylation by means of HPTLC (see above). The results of the exposure of Neu5,9Ac_2 and Neu5,4 Ac_2 with the EHEC *O*-acetylesterases are shown in Fig. 1A and Fig. 1B, respectively.

Incubation of Neu5,9Ac_2 with NanS-His (Fig. 1A, lane 4), NanS-p1-His (Fig. 1A, lane 5), NanS-p2-His (Fig. 1A, lane 6) and NanS-p4-His (Fig. 1A, lane 7) resulted in a HPTLC band that separated at the position of Neu5Ac reference in the chromatogram (Fig. 1A), indicating the de-*O*-acetylation of Neu5,9Ac_2 to Neu5Ac. Thus, all tested enzymes were capable to cleave the *O*-acetyl group from carbon atom C9, while Neu5,9Ac_2 remained intact in the buffer controls without enzyme (Fig. 1A, lane 3) when compared to the applied Neu5,9Ac_2 standard (Fig. 1A, lane 2). It should be mentioned that the Neu5,9Ac_2 preparation used in this study contained a small amount of Neu5Ac, which did not affect complete de-*O*-acetylation of the Neu5,9Ac_2 substrate to the Neu5Ac end product by the four EHEC *O*-acetylesterases. On the other hand, incubation of Neu5,4 Ac_2 with NanS-His (Fig. 1B, lane 4), NanS-p1-His (Fig. 1B, lane 5), NanS-p2-His (Fig. 1B, lane 6) and NanS-p4-His (Fig. 1B, lane 7) revealed bands at the position of Neu5,4 Ac_2 (Fig. 1B, lane 2), indicating that none of the recombinant *O*-acetylesterases was able to cleave the *O*-acetyl group from carbon C4.

3.2. NanoESI MS analysis of O-acetylesterase-treated BSM revealed enzymatic de-O-acetylation of mucin-bound mono-, di-, and tri-O-acetylated sialic acids

In a next step, we wanted to determine the specificity of recombinant NanS and NanS-p *O*-acetylesterases more precisely, since *O*-

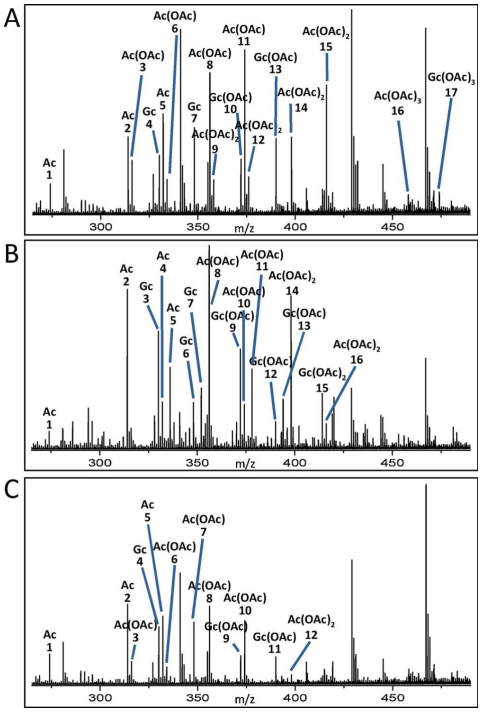

Fig. 2. nanoESI mass spectra of chemically released sialic acids from BSM (**A**), and of sialic acids obtained by chemical release upon pretreatment of BSM with NanS-His (**B**) and NanS-p1-His (**C**). (**A**) The overview mass spectrum of BSM-derived sialic acids indicates the presence of Neu5Ac, Neu5Gc, Neu5Ac(OAc), Neu5Gc(OAc), Neu5,9$(OAc)_2$, Neu5Ac$(OAc)_3$ and Neu5Gc$(OAc)_3$ in the *O*-glycans of BSM. (**B**) Ions corresponding to Neu5Ac, Neu5Gc, Neu5Ac(OAc), Neu5Gc(OAc), Neu5Ac$(OAc)_2$ and Neu5Gc$(OAc)_2$ were detected in the spectrum BSM-derived sialic acids after treatment with NanS-His. (**C**) The spectrum obtained from BSM-derived sialic acids after treatment with NanS-p1-His gives rise to ions corresponding to Neu5Ac, Neu5Gc, Neu5Ac(OAc), Neu5Gc(OAc) and Neu5Ac$(OAc)_2$. Ac, Neu5Ac; Ac(OAc), Neu5Ac(OAc); Ac$(OAc)_2$, Neu5Ac$(OAc)_2$; Ac$(OAc)_3$, Neu5Ac$(OAc)_3$. Gc, Neu5Gc; Gc(OAc), Neu5Gc(OAc); Gc$(OAc)_2$, Neu5Gc$(OAc)_2$; Gc$(OAc)_3$, Neu5Gc$(OAc)_3$. The numbered ions in the mass spectra depicted in panel A, B, and C, are listed in Table 2A–C, respectively.

acetyl groups do not only occur at positions C4 or C9, but also at C7 or C8 of sialic acids. Additionally, di- and tri-*O*-acetylated derivatives exist and unravelling NanS- or NanS-p-mediated de-*O*-acetylation of such sialic acids was also in the focus of our investigation. Consequently, we used BSM as a source for mucin-bound Neu5Ac, Neu5Gc and their diverse *O*-acetylated derivates, which are components of highly sialylated *O*-glycans of BSM as proposed in the manufacturer's data sheet (Merck, Darmstadt). Moreover, in recent work we could already demonstrate NanS-p activity in BSM (Saile et al., 2016).

In the initial experiment, BSM-derived sialic acids, obtained by mild acid treatment, were analyzed by nanoESI MS. The obtained mass spectrum demonstrated the presence of Neu5Ac as well as the corresponding mono-, di-, and tri-*O*-acetylated Neu5Ac derivatives (Fig. 2A). In addition, mono-, di-, and tri-*O*-acetylated Neu5Gc variants were detected. All detected sialic acid derivatives, which were chemically released from BSM *O*-glycans, are listed together with their *m/z* values

Table 2
Detected ions derived from chemically released sialic acids of BSM without enzyme (A), NanS-His treated BSM (B), and NanS-p1-His treated BSM (C), and corresponding *m/z* values.

No.	Sialic acid species	*m/z*	No.	Sialic acid species	*m/z*
A					
1	Neu5Ac$^+$ -H_2O	274.10	**10**	Neu5Gc(OAc)$^+$ -H + Na	372.09
2	Neu5Ac$^+$ -H + Na	314.09	**11**	Neu5Ac(OAc) [M + Na]$^+$	374.11
3	Neu5Ac(OAc)$^+$ -H_2O	316.11	**12**	Neu5Ac(OAc)$_2$$^+$	376.13
4	Neu5Gc$^+$ -H + Na	330.08	**13**	Neu5Gc(OAc) [M + Na]$^+$	390.10
5	Neu5Ac [M + Na]$^+$	332.10	**14**	Neu5Ac(OAc)$_2$$^+$ -H + Na	398.11
6	Neu5Ac(OAc)$^+$	334.12	**15**	Neu5Ac(OAc)$_2$ [M + Na]$^+$	416.12
7	Neu5Gc [M + Na]$^+$	348.09	**16**	Neu5Ac(OAc)$_3$ [M + Na]$^+$	458.13
8	Neu5Ac(OAc)$^+$ -H + Na	356.10	**17**	Neu5Gc(OAc)$_3$ [M + Na]$^+$	474.13
9	Neu5Ac(OAc)$_2$$^+$ -H_2O	358.11			
B					
1	Neu5Ac$^+$ -H_2O	274.10	**9**	Neu5Gc(OAc)$^+$ -H + Na	372.09
2	Neu5Ac$^+$ -H + Na	314.09	**10**	Neu5Ac(OAc) [M + Na]$^+$	374.11
3	Neu5Gc$^+$ -H + Na	330.08	**11**	Neu5Ac(OAc)$^+$ -2H + 2Na	378.08
4	Neu5Ac [M + Na]$^+$	332.10	**12**	Neu5Gc(OAc) [M + Na]$^+$	390.10
5	Neu5Ac$^+$ -2H + 2Na	336.07	**13**	Neu5Gc(OAc)$^+$ -2H + 2Na	394.07
6	Neu5Gc [M + Na]$^+$	348.09	**14**	Neu5Ac(OAc)$_2$$^+$ -H + Na	398.11
7	Neu5Gc$^+$ -2H + 2Na	352.06	**15**	Neu5Gc(OAc)$_2$$^+$ -H + Na	414.09
8	Neu5Ac(OAc)$^+$ -H + Na	356.10	**16**	Neu5Ac(OAc)$_2$ [M + Na]$^+$	416.12
C					
1	Neu5Ac$^+$ -H_2O	274.10	**7**	Neu5Gc [M + Na]$^+$	348.09
2	Neu5Ac$^+$ -H + Na	314.09	**8**	Neu5Ac(OAc)$^+$ -H + Na	356.10
3	Neu5Ac(OAc)$^+$ -H_2O	316.11	**9**	Neu5Gc(OAc)$^+$ -H + Na	372.09
4	Neu5Gc$^+$ -H + Na	330.08	**10**	Neu5Ac(OAc) [M + Na]$^+$	374.11
5	Neu5Ac [M + Na]$^+$	332.10	**11**	Neu5Gc(OAc) [M + Na]$^+$	390.10
6	Neu5Ac(OAc)$^+$	334.12	**12**	Neu5Ac(OAc)$_2$$^+$ -H + Na	398.11

in Table 2A.

In order to clarify, whether NanS-His and NanS-p-His *O*-acetylesterases are able to deacetylate mono-, di-, and tri-*O*-acetylated Neu5Ac or Neu5Gc species detected in BSM, we incubated BSM with either NanS-His, NanS-p1-His, NanS-p2-His or NanS-p4-His. Afterwards, the chemically released sialic acids were identified by nanoESI MS. Mass spectra of sialic acids obtained after exposure of BSM to the recombinant *O*-acetylesterases are exemplarily shown for NanS-His and NanS-p1-His in Fig. 2B and 2C, respectively. Identified sialic acids with the corresponding ion species and their *m/z* values are listed in Table 2B and C, respectively. Spectra of BSM-derived sialic acids upon exposure to NanS-His or NanS-p1-His evidenced predominant Neu5Ac and Neu5Gc core structures, accompanied by weaker signals, which could be assigned to mono-*O*-acetylated Neu5Ac(OAc) and Neu5Gc(OAc) species and the di-*O*-acetylated Neu5Ac(OAc)$_2$ derivative. Importantly, Neu5Ac(OAc)$_3$ and Neu5Gc(OAc)$_3$ variants, which appear in the spectra at *m/z* 458.13 and *m/z* 474.13, respectively, were undetectable after *O*-acetylesterase treatment. Similar results were obtained with recombinant NanS-p2-His and NanS-p4-His (data not shown). Our results show that NanS as well as NanS-p are able to release acetyl residues not only from mono-*O*-acetylated sialic acids, but also from di- and tri-*O*-acetylated sialic acids as demonstrated for Neu5Ac(OAc)$_2$, Neu5Ac(OAc)$_3$, and their di- and tri-*O*-acetylated Neu5Gc counterparts. Furthermore, we could show that NanS and NanS-p have the ability to deacetylate *O*-acetylated sialic acids in soluble form and linked to *O*-glycans of mammalian mucin.

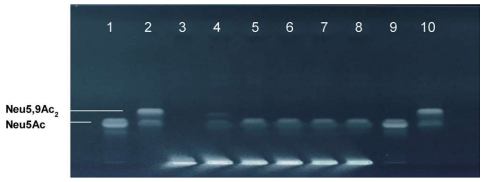

Fig. 3. Enzymatic cleavage of BSM-bound sialic acids by BTSA-His alone and combined with EHEC *O*-acetylesterases. Aliquots from enzymatic digests and standards were applied and sialic acids were stained with the DAPA reagent after chromatography. Approaches shown are BSM without enzymes (lane 3) and BSM with BTSA-His alone (lane 4), BTSA-His + NanS-His (lane 5), BTSA-His + NanS-p1-His (lane 6), BTSA-His + NanS-p2-His (lane 7), and BTSA-His + NanS-p4-His (lane 8). Neu5Ac (lanes 1, 9), and Neu5,9Ac$_2$ (lanes 2, 10) served as standards.

3.3. De-O-acetylation of mucin-bound sialic acids by NanS-His/NanS-p-His results in Neu5Ac and Neu5Gc substrates accessible to BTSA-His

To explore the general accessibility of mucin-derived Neu5Ac or Neu5Gc as a substrate for commensal sialidases like BTSA from *B. thetaiotaomicron* VPI-5482, we exposed BTSA-His and NanS-His or its homologs simultaneously to BSM, suggesting that de-*O*-acetylated sialic acids are convenient substrates for the sialidase. For this purpose, we incubated BSM with either sialidase BTSA-His, or coincubated BSM with sialidase BTSA-His and either NanS-His, NanS-p1-His, NanS-p2-His or NanS-p4-His. After an exposure time of 4 h, we determined the sialic acids released from BSM first by means of HPTLC. The results of the enzymatic digestions of BSM using BTSA-His alone and combined with NanS-His, NanS-p1-His, NanS-p2-His or NanS-p4-His are shown in Fig. 3. One mg/mL Neu5Ac (Fig. 3, lanes 1, 9) or Neu5,9Ac$_2$ (Fig. 3, lanes 2, 10) were used as standards. Due to its higher polarity, Neu5,9Ac$_2$ clearly separates beyond Neu5Ac allowing for unequivocal distinction of the two compounds. BSM incubated in buffer only without enzymes served as negative control (Fig. 3, lane 3).

The HPTLC analysis of enzymatically released sialic acids from BSM after incubation with BTSA-His alone (Fig. 3, lane 4) revealed two clearly DAPA-stained bands indicating the presence of enzymatically cleaved de-*O*-acetylated and mono-*O*-acetylated sialic acids. On the other hand, the combination of BTSA-His with NanS-His (Fig. 3, lane 5), NanS-p1-His (Fig. 3, lane 6), NanS-p2-His (Fig. 3, lane 7) or NanS-p4-His (Fig. 3, lane 8) revealed a single strong band, which separates at the position of Neu5Ac in the chromatogram as deduced from standard Neu5Ac (Fig. 3, lanes 1, 9).

Collectively, our experimental data gave evidence that BTSA-His is able to cleave Neu5Ac and Neu5,9Ac$_2$ from BSM. The enzyme seems not to exhibit any specificity for Neu5Ac or the *O*-acetylated derivative. Co-incubation of BSM with BTSA-His and *O*-acetylesterases revealed Neu5Ac only, while Neu5,9Ac$_2$ could not be detected. Due to the suggested capability of NanS-His and its homologs to deacetylate mono-, di-, and tri-*O*-acetylated sialic acids, the above described digests were repeated and the released sialic acids were detected using mass spectrometry. For this purpose, proteins of the digests were precipitated with methanol and the supernatants were directly applied to nanoESI MS. Mass spectra of sialic acids obtained from digests of BSM using BTSA-His alone, BTSA-His combined with NanS-His and BTSA-His combined with NanS-p1-His are shown in Fig. 4A, 4B and 4C, respectively. Identified sialic acids with the *m/z* values of the corresponding ions are listed in Table 3A–C, respectively. The mass spectrum obtained from BSM exposed to BTSA-His alone (Fig. 4A) shows ions derived from Neu5Ac and Neu5Gc basic structures and Neu5Ac(OAc) and Neu5Gc (OAc) (listed in Table 3A). Thus, BTSA-His was capable to cleave off Neu5Ac and Neu5Gc as well as the mono-*O*-acetylated counterparts from glycans. Di- and tri-*O*-acetylated sialic acids were not detected and

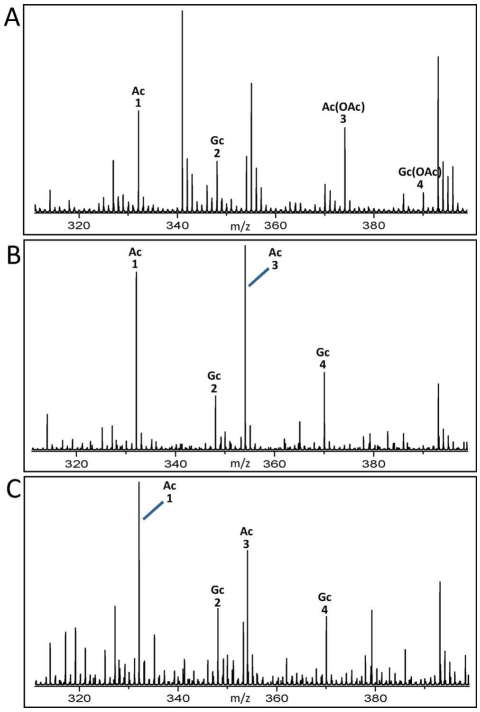

Fig. 4. nanoESI mass spectra of sialic acids obtained from digests of BSM treated with BTSA-His alone (A), BTSA-His combined with NanS-His (B), and BTSA-His combined with NanS-p1-His (C). The nanoESI MS analysis revealed ions which could be assigned to Neu5Ac(OAc) and Neu5Gc(OAc) only in the digests when using BTSA-His alone (A), whereas the combined digests (B and C) solely revealed the Neu5Ac and Neu5Gc cores. Ac, Neu5Ac; Ac(OAc), Neu5Ac (OAc); Gc, Neu5Gc; Gc(OAc), Neu5Gc(OAc). The numbered ions in the mass spectra depicted in panel A, B, and C, are listed in Table 3A–C, respectively.

Table 3
Detected ions assigned to sialic acids obtained by treatment of BSM with BTSA-His alone (A), BTSA-His combined with NanS-His (B), and BTSA-His combined with NanS-p1-His (C) and corresponding *m/z* values.

No.	Sialic acid species	*m/z*
A		
1	Neu5Ac $[M + Na]^+$	332.10
2	Neu5Gc $[M + Na]^+$	348.09
3	Neu5Ac(OAc) $[M + Na]^+$	374.11
4	Neu5Gc(OAc) $[M + Na]^+$	390.10
B		
1	Neu5Ac $[M + Na]^+$	332.10
2	Neu5Gc $[M + Na]^+$	348.09
3	Neu5Ac $[M\text{-}H + 2Na]^+$	354.08
4	Neu5Gc $[M\text{-}H + 2Na]^+$	370.07
C		
1	Neu5Ac $[M + Na]^+$	332.10
2	Neu5Gc $[M + Na]^+$	348.09
3	Neu5Ac $[M\text{-}H + 2Na]^+$	354.08
4	Neu5Gc $[M\text{-}H + 2Na]^+$	370.07

hence not released by BTSA-His. The combination of BTSA-His with the *O*-acetylesterases NanS-His or NanS-p1-His revealed ions indicating liberated Neu5Ac and Neu5Gc as shown in the mass spectra of Fig. 4B (listed in Table 3B) and Fig. 4C (listed in Table 3C), respectively. Both combined approaches revealed Neu5Ac and Neu5Gc but no *O*-acetylated sialic acids. The combination of BTSA-His and NanS-p2-His or NanS-p4-His gave the same results (data not shown).

In conclusion, the results demonstrate that the combination of both BTSA-His with NanS-His or NanS-p1-His, respectively, enhances the release of de-*O*-acetylated core Neu5Ac and Neu5Gc from mammalian mucin *O*-glycans, which might serve as a carbon and/or nitrogen source for EHEC.

4. Recombinant BTSA supports the growth of EHEC O157:H7 strain EDL933 on BSM dependent on the presence of NanS-p

We compared the growth behaviour (OD_{600} values) of EHEC O157:H7 strain EDL933 (NZ_CP015855) (Saile et al., 2016), EDL933Δ*nanS,* and EDL933Δ*nanS*Δ*nanS*-p1a-p7 for 24 h in growth medium containing BSM with or without BTSA-His. Despite an expected trend of the OD_{600} values, which decreased from 0.88, 0.8 to 0.76 for EDL933, EDL933Δ*nanS* and EDL933Δ*nanS*Δ*nanS*-p1a-p7, respectively, the differences were not statistically significant (Fig. 7). However, if BTSA-His was added to the medium, significant differences in the OD_{600} between EDL933Δ*nanS*Δ*nanS*-p1a-p7 and EDL933 or EDL933Δ*nanS* was detected with values of 0.91, 1.32 or 1.24, respectively (Fig. 5). For all strains, an increased OD_{600} was observed when BTSA-His was added, compared to the cultivation without BTSA-His. An explanation for this phenomenon is that both, Neu5Ac and Neu5,9Ac_2, terminally bound to the mucin, were released by BTSA-His, whereby the degradation of Neu5Ac was unaffected by the deletions. Thus, Neu5,9Ac_2 may be additionally utilized by EDL933 or EDL933Δ*nanS*. As a control a culture of EDL933Δ*nanS*Δ*nanS*-p1a-p7 was supplemented with NanS-p4-His, resulting in an OD_{600} comparable to that of a wild-type EDL933 culture.

The results of this study showed for the first time, that NanS-p *O*-acetylesterases are active in a mucin environment and can de-*O*-

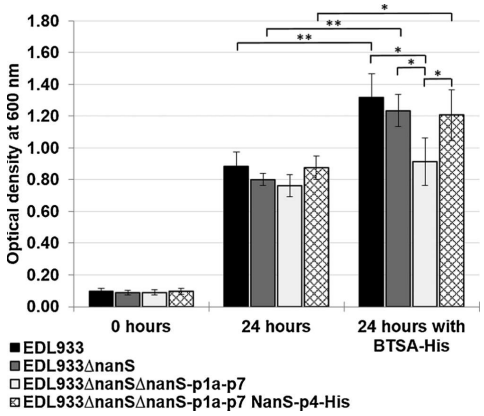

Fig. 5. Growth behaviour (OD_{600}) of *E. coli* O157:H7 strain EDL933 (black), EDL933Δ*nanS* (dark grey), EDL933Δ*nanS*Δ*nanS*-p1a-p7 (light grey) and EDL933Δ*nanS*Δ*nanS*-p1a-p7 supplemented with NanS-p4-His (hachured) in M9 minimal medium supplemented with 0.8% BSM at time points 0 and 24 h with or without BTSA-His. Error bars represent the standard deviations of 3 independent biological replicates. Significance was assessed with the paired Student *t* test (* $p < 0.05$; ** $p < 0.01$).

acetylate mucin-derived sialic acids *in vitro,* in combination with sialidase treatment.

5. Discussion

According to current concepts, EHEC have to outcompete the gut microbiota for access to energy sources for a successful infection process. Sialic acids are highly abundant in the large intestine, where they decorate the mucus layer(s) in terminally exposed positions on the mucin *O*-glycans. Sialic acids can be modified by up to three acetic acid esters at the positions of the hydroxyl groups of the sialic acid core including Neu5Gc species, whereby Neu5Gc and its derivatives do exist in all mammals except humans (Robbe et al., 2003; Varki and Schauer, 2009; Vimr, 2013). While pure Neu5Ac and Neu5Gc are readily cleavable by bacterial sialidases, *O*-acetylation protects them from enzymatic hydrolysis (Juge et al., 2016). The intestinal *E. coli* strains are generally considered competitors for EHEC strains owing to their similar genetic background. Both groups harbor genes for sialic acid metabolism and are able to utilize free sialic acids as carbon and nitrogen sources. Compared to commensal nonpathogenic *E. coli*, EHEC harbor several phage-encoded NanS homologs (NanS-p) (Saile et al., 2016, 2018). These enzymes catalyze de-*O*-acetylation of Neu5,9Ac_2 to Neu5Ac, which after sialidase action can be taken up and metabolized by both, commensal and pathogenic *E. coli* (Saile et al., 2016, 2018). Currently, it is not known, why EHEC strains contain a chromosomal *nanS* gene and multiple prophage-associated *nanS*-p genes. One hypothesis is that multiple *nanS-p* genes confer a gene dose effect to ensure substrate utilization and to outcompete commensal *E. coli* in the gut (Saile et al., 2016).

In this study, we demonstrated that recombinant NanS and NanS-p *O*-acetylesterases of EHEC O157:H7 strain EDL933 were able to de-*O*-acetylate sialic acids using BSM as a source for glycan-bound sialic acids. NanoESI MS spectra of chemically released sialic acids from BSM revealed ion signals of mono-, di-, and tri-*O*-acetylated Neu5Ac and Neu5Gc derivatives. NanS and NanS-p were capable to de-*O*-acetylate such *O*-acetylated sialic acids of BSM as determined by nanoESI MS analysis. Since various of the known microbial 9-*O*-acetylesterases are unable to de-*O*-acetylate Neu5,4 Ac_2, we probed the specificity of NanS and the NanS-p esterases with respect to cleavage of the acetyl ester at C4 of Neu5Ac. HPTLC analysis revealed that neither NanS nor the NanS-p esterases are able to split off the *O*-acetyl group from this position. Mass spectra of NanS- or NanS-p-treated mucin still exhibited ion signals of *O*-acetylated sialic acids. Since *O*-acetyl groups are mostly located at position C4 and/or C9, it is tempting to speculate that the detected mono-*O*-acetylated Neu5Ac-species represent the Neu5,4 Ac_2 isomer. Otherwise, an incomplete digestion or inactivation of the enzymes due to long-lasting incubation at elevated temperature could be the reason.

Furthermore, we could show that NanS and the NanS-p esterases deacetylate not only free sialic acids, but also glycosidically bound sialic acid of BSM. Although EHEC does not encode a sialidase, de-*O*-acetylation of bound sialic acids leads to liberation of acetate, which in turn can be readily used as an additional carbon source. However, de-*O*-acetylation results in sialidase-accessible Neu5Ac or Neu5Gc basic structures, which are common substrates of commensal sialidases, *e.g.*, BTSA of *B. thetaiotaomicron*. In order to verify this hypothesis, we probed the removal of BSM-bound sialic acids with BTSA-His. This sialidase hydrolyzes pure Neu5Ac and Neu5Gc and also the mono-*O*-acetylated counterparts Neu5Ac(OAc) and Neu5Gc(OAc). Di- and tri-*O*-acetylated sialic acids were resistant towards BTSA giving evidence for a protective effect of *O*-acetylated sialic acids from degradation.

Combined exposure of BTSA with NanS or NanS-p resulted in release of Neu5Ac and Neu5Gc from BSM, which can then be metabolized in soluble form by EHEC. Even though *B. thetaiotaomicron* encodes sialidases, this bacterium lacks the *nan* operon and is unable to utilize released sialic acids as a nutrient. Therefore, the complementary action of both enzymes may result in a higher content of free sialic acids in the gut. In order to assess the growth-promoting effect of a mixed enzyme treatment, we cultivated the EHEC O157:H7 wildtype strain EDL933, and its mutant strains EDL933Δ*nanS* and EDL933Δ*nanS*Δ*nanS*-p1a-p7 on mucin, mucin supplemented with BTSA-His and mucin exposed to BTSA-His combined with NanS-p-His. The combination revealed a significantly enhanced bacterial growth suggesting that EHEC exploits sialidases of commensal gut bacteria, and is supported by own *O*-acetylesterases to get access to *O*-glycan-derived sialic acids of intestinal mucus. On principle, sialic acids *per se* and particularly the *O*-acetylated species prevent from subsequent digest of desialylated glycans by glycosidases (Phansopa et al., 2015). Collectively, the sialidase-mediated release of sialic acids combined with complementary saponification of acetic acid side chains allows for ensuing stepwise digest of desialylated glycans by glycosidases such as galactosidases, *N*-acetylhexosaminidases or fucosidases. It can be speculated that the primary role of NanS-p esterases during infection could be the perturbation of the mucus layer integrity and initiation of commensal degradation, which may enable or at least facilitate EHEC strains crossing of the mucus barrier and getting access to the epithelial cells in the colon.

Interestingly, a number of viruses do exist, *e.g.*, human coronavirus OC43 or mouse hepatitis virus, encoding for *O*-acetylesterases that may play a comparative role like the bacterial esterases during infection. These enzymes are mostly specific for Neu5,9Ac_2 (human coronavirus OC43) or Neu5,4 Ac_2 (mouse hepatitis virus) (Strasser et al., 2004). The influenza C virus encodes for an *O*-acetylesterase being capable to hydrolyze glycosidically bound Neu5,9Ac_2, Neu5,7(8),9Ac_3 and Neu5Gc9Ac of BSM in a similar manner like NanS-p-His (Strasser et al., 2004). This *O*-acetylesterase enables the virus to attach to the host membrane, destroys receptors accompanied with altered signal transduction in the host cell and enables viral entry (Ayora-Talavera, 2018; Matrosovich et al., 2015; Strasser et al., 2004).

We propose that NanS-p esterases might have an important complementary role during mixed viral and bacterial infections due to their broad specificity compared to that of the influenza C virus esterase. Thus, EHEC NanS-p esterases, which have the potency to de-*O*-acetylate mono-, di-, and tri-*O*-acetylated receptor-bound sialic acids, might have a yet underestimated functional role throughout initial infection and immune evasion of EHEC. Further work is necessary to clarify the role of NanS-p *O*-acetylesterases with regard to their involvement in maintenance and successful colonization of EHEC strains.

Acknowledgements

This work was supported by grant Schm1360/6-1 from the Deutsche Forschungsgemeinschaft (DFG). J. M. further acknowledges financial support by the German Center for Infection Research (DZIF, TTU 06.801). We thank Anja Voigt for experimental support in cloning of *nanS*.

References

Aperce, C.C., Heidenreich, J.M., Drouillard, J.S., 2014. Capacity of the bovine intestinal mucus and its components to support growth of *Escherichia coli* O157:H7. Animal 8, 731–737.
Ayora-Talavera, G., 2018. Sialic acid receptors: focus on their role in influenza infection. JRLCR 10, 1–11.
Barnett Foster, D., 2013. Modulation of the enterohemorrhagic *E. coli* virulence program through the human gastrointestinal tract. Virulence 4, 315–323.
Bertin, Y., Chaucheyras-Durand, F., Robbe-Masselot, C., Durand, A., de La Foye, A., Harel, J., Cohen, P.S., Conway, T., Forano, E., Martin, C., 2013. Carbohydrate utilization by enterohaemorrhagic *Escherichia coli* O157:H7 in bovine intestinal content. Environ. Microbiol. 15, 610–622.
Chang, D.-E., Smalley, D.J., Tucker, D.L., Leatham, M.P., Norris, W.E., Stevenson, S.J., Anderson, A.B., Grissom, J.E., Laux, D.C., Cohen, P.S., Conway, T., 2004. Carbon nutrition of *Escherichia coli* in the mouse intestine. Proc. Natl. Acad. Sci. U S A 101, 7427–7432.
Conway, T., Cohen, P.S., 2015. Commensal and pathogenic *Escherichia coli* metabolism in the gut. Microbiol. Spectr. 3.
Conway, T., Krogfelt, K.A., Cohen, P.S., 2004. The kife of commensal *Escherichia coli* in

the mammalian intestine. EcoSal Plus 1.
Corfield, A.P., Wagner, S.A., Clamp, J.R., Kriaris, M.S., Hoskins, L.C., 1992. Mucin degradation in the human colon: production of sialidase, sialate O-acetylesterase, N-acetylneuraminate lyase, arylesterase, and glycosulfatase activities by strains of fecal bacteria. Infect. Immun. 60, 3971–3978.
Corfield, A.P., Wagner, S.A., O'Donnell, L.J., Durdey, P., Mountford, R.A., Clamp, J.R., 1993. The roles of enteric bacterial sialidase, sialate O-acetyl esterase and glycosulfatase in the degradation of human colonic mucin. Glycoconj. J. 10, 72–81.
Erdem, A.L., Avelino, F., Xicohtencatl-Cortes, J., Girón, J.A., 2007. Host protein binding and adhesive properties of H6 and H7 flagella of attaching and effacing Escherichia coli. J. Bacteriol. 189, 7426–7435.
Fabich, A.J., Jones, S.A., Chowdhury, F.Z., Cernosek, A., Anderson, A., Smalley, D., McHargue, J.W., Hightower, G.A., Smith, J.T., Autieri, S.M., Leatham, M.P., Lins, J.J., Allen, R.L., Laux, D.C., Cohen, P.S., Conway, T., 2008. Comparison of carbon nutrition for pathogenic and commensal *Escherichia coli* strains in the mouse intestine. Infect. Immun. 76, 1143–1152.
Fox, J.T., Drouillard, J.S., Shi, X., Nagaraja, T.G., 2009. Effects of mucin and its carbohydrate constituents on *Escherichia coli* O157 growth in batch culture fermentations with ruminal or fecal microbial inoculum. J. Anim. Sci. 87, 1304–1313.
Hews, C.L., Tran, S.-L., Wegmann, U., Brett, B., Walsham, A.D.S., Kavanaugh, D., Ward, N.J., Juge, N., Schüller, S., 2017. The StcE metalloprotease of enterohaemorrhagic *Escherichia coli* reduces the inner mucus layer and promotes adherence to human colonic epithelium ex vivo. Cell Microbiol. 19.
Johansson, M.E.V., Thomsson, K.A., Hansson, G.C., 2009. Proteomic analyses of the two mucus layers of the colon barrier reveal that their main component, the Muc2 mucin, is strongly bound to the Fcgbp protein. J. Proteome Res. 8, 3549–3557.
Johansson, M.E.V., Larsson, J.M.H., Hansson, G.C., 2011. The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. Proc. Natl. Acad. Sci. U. S. A. 108, 4659–4665.
Juge, N., Tailford, L., Owen, C.D., 2016. Sialidases from gut bacteria: a mini-review. Biochem. Soc. Trans. 44, 166–175.
Marcobal, A., Southwick, A.M., Earle, K.A., Sonnenburg, J.L., 2013. A refined palate: bacterial consumption of host glycans in the gut. Glycobiology 23, 1038–1046.
Matrosovich, M., Herrler, G., Klenk, H.D., 2015. Sialic acid receptors of viruses. Top. Curr. Chem. 367, 1–28.
Nataro, J.P., Kaper, J.B., 1998. Diarrheagenic Escherichia coli. Clin. Microbiol. Rev. 11, 142–201.
Nguyen, Y., Sperandio, V., 2012. Enterohemorrhagic *E. coli* (EHEC) pathogenesis. Front Cell Infect. Microbiol. 2, 90.
Nübling, S., Eisele, T., Stöber, H., Funk, J., Polzin, S., Fischer, L., Schmidt, H., 2014. Bacteriophage 933W encodes a functional esterase downstream of the Shiga toxin 2a operon. Int. J. Med. Microbiol. 304, 269–274.
O'Brien, A.D., Melton, A.R., Schmitt, C.K., McKee, M.L., Batts, M.L., Griffin, D.E., 1993. Profile of *Escherichia coli* O157:H7 pathogen responsible for hamburger-borne outbreak of hemorrhagic colitis and hemolytic uremic syndrome in Washington. J. Clin. Microbiol. 31, 2799–2801.
Page, A.V., Liles, W.C., 2013. Enterohemorrhagic *Escherichia coli* infections and the hemolytic-uremic syndrome. Med. Clin. North Am. 97, 681–695 xi.
Park, K.-H., Kim, M.-G., Ahn, H.-J., Lee, D.-H., Kim, J.-H., Kim, Y.-W., Woo, E.-J., 2013. Structural and biochemical characterization of the broad substrate specificity of *Bacteroides thetaiotaomicron* commensal sialidase. Biochim. Biophys. Acta 1834, 1510–1519.
Peri, S., Kulkarni, A., Feyertag, F., Berninsone, P.M., Alvarez-Ponce, D., 2018. Phylogenetic distribution of CMP-Neu5Ac hydroxylase (CMAH), the enzyme synthetizing the proinflammatory human xenoantigen Neu5Gc. Genome Biol. Evol. 10, 207–219.
Phansopa, C., Kozak, R.P., Liew, L.P., Frey, A.M., Farmilo, T., Parker, J.L., Kelly, D.J., Emery, R.J., Thomson, R.I., Royle, L., Gardner, R.A., Spencer, D.I.R., Stafford, G.P., 2015. Characterization of a sialate-*O*-acetylesterase (NanS) from the oral pathogen Tannerella forsythia that enhances sialic acid release by NanH, its cognate sialidase. Biochem. J. 472, 157–167.
Rangarajan, E.S., Ruane, K.M., Proteau, A., Schrag, J.D., Valladares, R., Gonzalez, C.F., Gilbert, M., Yakunin, A.F., Cygler, M., 2011. Structural and enzymatic characterization of NanS (YjhS), a 9-O-Acetyl N-acetylneuraminic acid esterase from *Escherichia coli* O157:H7. Protein Sci. 20, 1208–1219.
Rangel, A., Steenbergen, S.M., Vimr, E.R., 2016. Unexpected diversity of *Escherichia coli* sialate O-Acetyl esterase NanS. J. Bacteriol. 198, 2803–2809.
Riley, L.W., Remis, R.S., Helgerson, S.D., McGee, H.B., Wells, J.G., Davis, B.R., Hebert, R.J., Olcott, E.S., Johnson, L.M., Hargrett, N.T., Blake, P.A., Cohen, M.L., 1983. Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. N. Engl. J. Med. 308, 681–685.
Robbe, C., Capon, C., Maes, E., Rousset, M., Zweibaum, A., Zanetta, J.-P., Michalski, J.-C., 2003. Evidence of regio-specific glycosylation in human intestinal mucins: presence of an acidic gradient along the intestinal tract. J. Biol. Chem. 278, 46337–46348.
Robinson, L.S., Lewis, W.G., Lewis, A.L., 2017. The sialate *O*-acetylesterase EstA from gut Bacteroidetes species enables sialidase-mediated cross-species foraging of 9-*O*-acetylated sialoglycans. J. Biol. Chem. 292, 11861–11872.
Saile, N., Voigt, A., Kessler, S., Stressler, T., Klumpp, J., Fischer, L., Schmidt, H., 2016. *Escherichia coli* O157:H7 strain EDL933 harbors multiple functional prophage-associated genes necessary for the utilization of 5-*N*-acetyl-9-*O*-acetyl neuraminic acid as a growth substrate. Appl. Environ. Microbiol. 82, 5940–5950.
Saile, N., Schwarz, L., Eißenberger, K., Klumpp, J., Fricke, F.W., Schmidt, H., 2018. Growth advantage of *Escherichia coli* O104:H4 strains on 5-*N*-acetyl-9-*O*-acetyl neuraminic acid as a carbon source is dependent on heterogeneous phage-borne nanS-p esterases. Int. J. Med. Microbiol. 308, 459–468.
Sambrook, J., Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor (N.Y.).
Scheutz, F., Nielsen, E.M., Frimodt-Møller, J., Boisen, N., Morabito, S., Tozzoli, R., Nataro, J.P., Caprioli, A., 2011. Characteristics of the enteroaggregative shiga toxin/verotoxin-producing *Escherichia coli* O104:H4 strain causing the outbreak of haemolytic uraemic syndrome in Germany, May to June 2011. Euro Surveill. 16.
Stecher, B., Hardt, W.-D., 2011. Mechanisms controlling pathogen colonization of the gut. Curr. Opin. Microbiol. 14, 82–91.
Steenbergen, S.M., Jirik, J.L., Vimr, E.R., 2009. YjhS (NanS) is required for *Escherichia coli* to grow on 9-*O*-acetylated N-acetylneuraminic acid. J. Bacteriol. 191, 7134–7139.
Steil, D., Pohlentz, G., Legros, N., Mormann, M., Mellmann, A., Karch, H., Müthing, J., 2018. Combining mass spectrometry, surface acoustic wave interaction analysis, and cell viability assays for characterization of Shiga toxin subtypes of pathogenic *Escherichia coli* bacteria. Anal. Chem. 90, 8989–8997.
Stevens, M.P., Frankel, G.M., 2014. The locus of enterocyte effacement and associated virulence factors of enterohemorrhagic *Escherichia coli*. Microbiol. Spectr. 2 EHEC-0007-2013.
Strasser, P., Unger, U., Strobl, B., Vilas, U., Vlasak, R., 2004. Recombinant viral sialate-O-acetylesterases. Glycoconj. J. 20, 551–561.
Studier, F.W., Moffatt, B.A., 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. 189, 113–130.
Tailford, L.E., Owen, C.D., Walshaw, J., Crost, E.H., Hardy-Goddard, J., Le Gall, G., de Vos, W.M., Taylor, G.L., Juge, N., 2015. Discovery of intramolecular trans-sialidases in human gut microbiota suggests novel mechanisms of mucosal adaptation. Nat. Commun. 6, 7624.
Torres, A.G., Zhou, X., Kaper, J.B., 2005. Adherence of diarrheagenic *Escherichia coli* strains to epithelial cells. Infect. Immun. 73, 18–29.
Varki, A., Schauer, R., 2009. Essentials of Glycobiology: Sialic Acids, 2nd ed. Cold Spring Harbor (NY).
Vimr, E.R., 2013. Unified theory of bacterial sialometabolism: how and why bacteria metabolize host sialic acids. ISRN Microbiol. (2013), 816713.
Vimr, E.R., Kalivoda, K.A., Deszo, E.L., Steenbergen, S.M., 2004. Diversity of microbial sialic acid metabolism. Microbiol. Mol. Biol. Rev. 68, 132–153.

5. Abschlussdiskussion

Pathogene *E. coli,* im Besonderen EHEC, tragen mehrere *nanS*-homologe offene Leserahmen in Prophagen lokalisiert in ihren Chromosomen (Vimr, 2013). Im Rahmen dieser Arbeit wurden sie als *nanS*-p deklariert und mit jeweils einer Zahl versehen (Kapitel 2). Aufgrund ihrer Nukleotidsequenzlängen unterteilte Vimr (2013) sie in kurz, lang, fragmentiert, mittellang, partiell oder sehr lang. In früheren Studien wurde gezeigt, dass NanS-p1 von *E. coli* O157:H7 Stamm EDL933 als *O*-Acetyl Esterase aktiv ist (Nübling et al., 2014) und unter Dickdarm-simulierenden Bedingungen gebildet wird (Polzin et al., 2013).
Eine wichtige Frage dieser Arbeit war, ob auch andere *nanS*-p Gene, neben *nanS*-p1, für funktionelle Proteine kodieren und ob deren Aktivitäten zur Energiegewinnung der Bakterien beitragen. Außerdem sollte geklärt werden, warum EHEC und andere pathogene *E. coli* multiple *nanS*-p Gene in ihren Chromosomen haben. Eine weitere Frage war, ob pathogene *E. coli* durch die NanS-p Proteine wettbewerbsfähig gegenüber kommensalen *E. coli* sind, indem sie das Substratspektrum erweitern oder die effiziente Nutzung einer Nährstoffnische ermöglichen. Zur Beantwortung wurden genetische, funktionelle und mikrobiologische Analysen durchgeführt.

Das zentrale Ergebnis in dieser Dissertation ist, dass alle untersuchten rekombinant hergestellten NanS-p Proteine in den pathogenen *E. coli* Stämmen EDL933 (Kapitel 2) und LB226692 (Kapitel 3) funktionelle Esterasen mit ähnlichen Temperatur- und pH- Optima sind. *In silico* Analysen zeigten, dass die Aminosäuresequenzen die Domäne 303 und die unbekannte Domäne 1737 besitzen (Kapitel 2). Die Domäne 303 ist für die Sialinsäureesterasefunktion von NanS verantwortlich. Auch die NanS-p Esterasen gehören zu den Sialinsäureesterasen. Für NanS-p2 und NanS-p4 wurde gezeigt, dass sie Acetat aus chemisch synthetisiertem Neu5,9Ac_2 und Rinderspeicheldrüsenmuzin freisetzen (Kapitel 2). Konvertiertes Neu5Ac und Acetat könnten den Bakterien dann als Energiequellen zur Verfügung stehen.
Zur Überprüfung dieser Hypothese sollten Wachstumskurven mit *nanS*/*nanS*-p Deletionsmutantenstämmen der pathogenen Stämme EDL933 und C227-11φcu mit Neu5,9Ac_2 in einem Minimalmedium erstellt werden. Für die Deletionen wurden die

Nukleotidsequenzen des *nanS* Gens und der *nanS*-p Gene auf Grundlage der Chromosomensequenz AE005174 des EDL933 Stammes (Perna et al., 2001), die bei NCBI hinterlegt ist, ermittelt. Ein *nanS* Gen, sechs partielle und sieben lange *nanS*-p Gene konnten *in silico* nachgewiesen werden (Daten nicht publiziert).
Die Deletionsmutanten wurden mit einer Deletionsmethode von Datsenko & Wanner (2000) konstruiert. Dazu wurden Kanamycinresistenzkassetten (kurz: Kassetten) konstruiert, die ein Kanamycinresistenzgen mit flankierenden Flippase Erkennungsstellen (FRT Stellen) besitzen. Mit entsprechenden Oligonukleotidprimern wurden an den 5‘- und 3‘- Enden zum deletierenden Gen homologe Enden angebracht. Durch homologe Rekombination *in vivo* wurde das zu deletierende Gen gegen die Kassette im bakteriellen Chromosom ausgetauscht und dadurch deletiert. Positive Bakterienkolonien konnten anhand der Antibiotikaresistenz selektiert werden. Anschließend wurde das Plasmid pCP20 in die Zellen transformiert, dass eine Flippase kodiert. Die Flippase schnitt das Antibiotikaresistenzgen an den FRT Stellen aus (Cherepanov & Wackernagel, 1995). Im Chromosom blieb eine FRT Stelle als „Narbe“ zurück. Als erstes sollte *nanS* und dann sukzessive alle langen *nanS*-p Gene deletiert werden. Während *nanS* einfach deletiert werden konnte, gab es zunächst Komplikationen bei den Deletionen der *nanS*-p Gene. Aufgrund der Homologien der *nanS*-p Gene zueinander war eine zielgerichtete Deletion nicht möglich, da die 5‘- und 3‘-Enden der Kassette meist zu mehreren *nanS*-p Genen homolog waren. Außerdem war ein Fehleinbau der Kassette in die Narben im Chromosom aufgrund homologer Bereiche der Kassette nicht zu vermeiden. Durch Selektion mit Kanamycin-haltigen Agarböden wuchsen also nicht nur Kolonien, deren Bakterien die Deletion des gewünschten Gens aufwiesen, sondern auch Kolonien, deren Bakterien ein anderes *nanS*-p Gen deletiert hatten, oder bei den die Kassette wieder in eine Narbe eingebaut wurde. Je mehr Gene deletiert waren, desto mehr Narben waren im Chromosom vorhanden und desto mehr falsch positive Bakterienkolonien wuchsen. Mittels Polymerasekettenreaktion (PCR) wurden anschließend positive Deletionsmutanten identifiziert. Im Laufe der Deletionsarbeiten stellte sich jedoch heraus, dass einzelne *nanS*-p Gene mittels PCR im Wildtyp nicht nachweisbar waren oder die nachgewiesenen *nanS*-p Gene nach Sequenzierung eine andere Nukleotidsequenz aufwiesen, wie erwartet wurde. Die Chromosomensequenz

AE005174 stimmte mit der Sequenz des in dieser Arbeit verwendeten EDL933 also nicht überein. Daher mussten die Arbeiten zunächst komplett eingestellt und die vorhandenen Deletionsmutanten konserviert werden.

Das Genom des EDL933 Stammes wurde zur Sequenzierung isoliert. An der ETH Zürich wurde daraus eine zusammenhängende Nukleotidsequenz (CP015855) erstellt (Kapitel 2). Parallel wurde von Latif et al. (2014) eine weitere EDL933 Genomsequenz (NZ_CP008957) veröffentlicht. *In silico* Analysen zur Quantität der *nanS*-p Gene in NZ_CP008957 und CP015855 ergaben zehn bzw. sieben Prophagen-kodierte lange *nanS*-p Gene (Kapitel 2). Sechs *nanS*-p Nukleotidsequenzen der Sequenz CP015855 sind auch in NZ_CP008957 identisch vorhanden. Eine *nanS*-p Nukleotidsequenz (*nanS*-p1a) aus CP015855 unterscheidet sich von der entsprechenden Sequenz (*nanS*-p1) in NZ_CP008957. Ein *nanS* Gen ist in beiden Sequenzen, wie erwartet vorhanden, jedoch konnten keine partiellen *nanS*-p Gene nachgewiesen werden. Die partiellen *nanS*-p Nukleotidsequenzen und vermutlich auch die fragmentierten, mittellangen und sehr langen Sequenzen, wie sie von Vimr (2013) benannt wurden, scheinen vielmehr Artefakte aus früheren Sequenzierungen zu sein. Sie entstanden vermutlich durch fehlerhafte Überlagerungen von homologen Nukleotidsequenzen der Prophagen. Durch den modernen Einsatz neuester Next-Generation Sequenzierungstechniken ist die Fehleranfälligkeit heutzutage sehr viel geringer. Die Unterschiede der beiden Sequenzen NZ_CP008957 und CP015855 lassen sich durch Deletionen und Insertionen in den Chromosomen unter flexiblen Lagerungsbedingungen der Bakterienkonserven, im Laufe der Zeit, erklären (Papadopoulos et al., 1999; Fellner et al., 2016).

Die konservierten EDL933 Deletionsmutantenstämme (siehe oben) wurden auf Grundlage der neuen Chromosomensequenz CP015855 auf ihre vorhandenen *nanS*/*nanS*-p Deletionen im Chromosom überprüft. Die Identitäten der bereits deletierten und der noch vorhandenen *nanS*-p Gene konnten mittels PCR zweifelsfrei geklärt werden. Sukzessive wurden die restlichen *nanS*-p Gene erfolgreich deletiert.

Auch das Genom des *E. coli* O104:H4 Stammes LB226692 wurde im Laufe dieser Arbeit sequenziert und entsprechend bioinformatisch analysiert. Überraschend ist, dass im LB226692 Chromosom kein *nanCMS* Operon vorhanden ist (Kapitel 3). Es

kommen jedoch fünf Prophagen-kodierte *nanS*-p Gene vor. In keiner der analysierten Genomsequenzen von Stämmen des Serotyps O104:H4 der NCBI Datenbank konnte *nanCMS* nachgewiesen werden (Kapitel 3). Deletionsmutanten wurden von einer stx_{2a}-Prophagen-negativen Varianten des O104:H4 C227-11 Isolates (C227-11φcu) konstruiert. Aufgrund des fehlenden stx_{2a}-Prophagens in C227-11φcu, der normalerweise NanS-p15 kodiert, sind im Chromosom dieser Varianten nur vier *nanS*-p Gene vorhanden.

In Kultivierungsexperimenten konnte gezeigt werden, dass die NanS-p *O*-Acetyl Esterasen tatsächlich für eine Energiegewinnung der Stämme aus Neu5,9Ac_2 genutzt werden (Kapitel 2 und 3). Erstmalig wurde darüber hinaus gezeigt, dass NanS für die Vermehrung von *E. coli* mit Neu5,9Ac_2 nicht essentiell ist solange NanS-p Proteine gebildet werden oder sich funktionelle NanS-p Proteine in der Zellumgebung befinden um Neu5Ac und/oder Acetat bereitzustellen (Kapitel 2 und 3). Etwa zeitgleich mit der Veröffentlichung von Saile et al. (2016) (Kapitel 2) publizierten Rangel et al. (2016) eine verwandte Studie, die die hier präsentierten Ergebnisse unterstützt. Auch hier konnte sich die *nanS*-Deletionsmutante von EDL933 uneingeschränkt, wie der Wildtyp, mit Neu5,9Ac_2 vermehren. Rangel et al. (2016) führten das auf die Anwesenheit von *nanS*-p Genen im Chromosom zurück, bewiesen dies experimentell aber nicht. In Kapitel 4, im Teil der Publikation, der von Stefanie Feuerbaum und Gottfried Pohlentz bearbeitet wurde, wird nach HPTLC und nanoESI Analysen in entsprechenden Experimenten eine bisher unbekannte Substratspezifität von NanS und NanS-p Proteinen detektiert. Neben Neu5,9Ac_2 und Neu5,8Ac_2 wurden auch andere zweifach und dreifach *O*-acetylierte Neu5Ac und Neu5Gc Derivate deacetyliert. Neu5,4Ac_2 ist jedoch kein Substrat von NanS oder NanS-p Proteinen. Im Menschen kommen Neu5Gc Derivate normalerweise nicht vor (Kapitel 1.7). In dieser Studie ist die Anwesenheit von Neu5Gc und dessen Derivate im Muzin auf seine Herkunft aus Rindern zurückzuführen.

Mit den hier durchgeführten Analysen konnte gezeigt werden, dass *E. coli* durch die Aktivitäten von NanS-p Proteinen lediglich Zugang zu denselben Substraten erhält, wie durch die Aktivität von NanS. Kodieren *E. coli* Bakterien jedoch nicht für NanS, wie die Isolate LB226692 oder C227-11φcu sind die NanS-p Proteine allein verantwortlich für die Energiegewinnung aus diesen Substraten. Der bakterielle Vorteil der

Anwesenheit von NanS-p Proteinen gegenüber der von NanS könnte die Lokalisierung der NanS-p Proteine, die Quantität oder eine Anpassung an verschiedene Habitate durch unterschiedliche Enzymcharakteristika sein.

Durch Analyse der Wachstumskurven von C227-11φcu, C227-11φcu∆*nanS*-p11, C227-11φcu∆*nanS*-p11,12, C227-11φcu∆*nanS*-p11,12,13 und C227-11φcu∆*nanS*-p11,12,13,14 in Minimalmedium mit Neu5,9Ac_2 als Kohlenhydratquelle konnte einen Gen-Dosis Effekt durch eine Verlängerung der Generationszeiten der Stämme mit abnehmender Anzahl an *nanS*-p Genen im Chromosom gezeigt werden (Kapitel 3). Im Rückkehrschluss bedeutet dies, zumindest in diesem Modell, dass eine *E. coli* Kultur mit Sialinsäurederivaten effizienter wachsen kann, je mehr *nanS*-p Gene im Chromosom des Stammes vorhanden sind. Es kann spekuliert werden, dass der Gen-Dosis Effekt für eine schnelle Vermehrung im Darm und die Wettbewerbsfähigkeit gegenüber der Darmmikrobiota essentiell ist.

Es wurden Co-Kultivierungsexperimente *in vitro* mit dem apathogenen Stamm AMC 198 und dem STEC Stamm C227-11φcu oder dessen Deletionsmutantenstämme mit Neu5,9Ac_2 als Substrat durchgeführt. Dabei sollte überprüft werden in wie weit die *nanS*-p Gene zur Wettbewerbsfähigkeit des C227-11φcu Stammes beitragen. C227-11φcu konnte im Verlauf der Kultivierung zu hohen Lebendkeimzahlen heranwachsen, während sich AMC 198 kaum vermehrte (Kapitel 3). Auch in der Co-Kultivierung der Deletionsmutanten C227-11φcu∆*nanS*-p11,12,13 und AMC 198 setzte sich die Mutante durch. Dieses Ergebnis unterstützt die Hypothese, dass NanS-p Proteine von pathogenen *E. coli* Stämmen genutzt werden, um sich gegen metabolisch ähnlich aktive kommensale *E. coli* Stämme durchzusetzen.

Sialinsäurederivate kommen im Dickdarm jedoch gebunden an Muzinen vor und *E. coli* selbst bildet keine Sialidasen, die zur Freisetzung der Sialinsäuren nötig wären. Gebundene Sialinsäuren scheinen deshalb zunächst nutzlos für *E. coli*. Ein Modell wie Muzin-gebundene Sialinsäuren für *E. coli* zugänglich werden könnten, wurde in Kapitel 4 vorgestellt. Die Interaktion von NanS-p Proteinen und einer Sialidase von anaeroben Mitgliedern der Darmmikrobiota ist dabei nötig. Die Sialinsäurederivate werden dabei zunächst durch NanS-p Proteine de-*O*-acetyliert, und anschließend durch die Sialidase von der Glykankette gespalten. *O*-Acetylierungen

verhindern nämlich eine Sialidaseaktivität (Juge et al., 2016). Aufgrund dieser Ergebnisse ist eine Beteiligung von Mitgliedern der Darmmikrobiota am EHEC Sialinsäurekatabolismus vorstellbar und Muzin könnte als Sialinsäurelieferant genutzt werden.

Die spezielle Lokalisation der *nanS*-p Gene zwischen dem Antiterminator Q Gen und den Genen der Lysekassette in den Prophagen könnte den Bakterien einen Vorteil bei der Energiegewinnung bringen und ist daher diskutabel. NanS-p Proteine könnten möglicherweise, wie Stx, durch Phagenlyse aus der Zelle sekretiert und direkt am Substrat aktiv werden, während sich NanS im Periplasma der Bakterienzelle befindet und daher eine Barriere gegenüber dem Substrat besitzt. Unterschiedliche Phasen im Wachstum der *E. coli* Population könnten einen Einfluss auf die Phagenlyse einzelner Bakterienzellen der Population haben. Dabei würde verhindert werden, dass alle Zellen der Population gleichzeitig lysieren und die Population ausstirbt (Imamovic et al., 2016). Smith (2001) stellte die „Cheater Hypothese“ auf, der zufolge Mutanten einer Population, die einen bestimmten Virulenzfaktor nicht bilden, von den Zellen der Population profitieren, die diesen Virulenzfaktor sekretieren. Dabei können die Mutanten ihre Energie gezielt für die Vermehrung einsetzen und das Gen des Virulenzfaktors anschließend durch horizontalen Gentransfer erwerben. In diesem Fall würden nicht-lysierte Zellen der Population direkt von bereitgestelltem Neu5Ac und/oder Acetat aus zwei- und dreifach *O*-acetylierten Derivaten profitieren. Obendrein würde Energie durch die Neu5Ac-vermittelte Herabregulierung der Transkription der Virulenzgene eingespart (Kapitel 1.6.2) (Le Bihan et al., 2015; Le Bihan et al., 2017). Bei dieser Hypothese würde es durch Energie- und Zeitersparnis bei der Vermehrung zu einem Kolonisationserfolg oder zumindest zur Co-Existenz mit kommensalen *E. coli* Stämmen im Dickdarm kommen. Im Gegensatz zur „Cheater Hypothese“ müsste das Gen des Virulenzfaktors jedoch nicht mehr erworben werden, denn der Prophage, der das NanS-p Protein kodiert, wäre bereits lysogen im Chromosom integriert. Jedoch würde eine Infektion von kommensalen *E. coli* durch freigesetzten Phagen diese zu NanS-p Spendern umwandeln und die pathogenen *E. coli* Bakterien würden als „Cheater“ erneut profitieren. Zumindest für Stx-Phagen, die kommensale *E. coli* zu STEC mutierten und

diese zu Stx-Spendern konvertieren, wurde dies mehrfach gezeigt (Gamage et al., 2004; Iversen et al., 2015; Xiaoli et al., 2018).
Völlig unberücksichtigt blieb bisher, dass eine Abspaltung endständiger Sialinsäuren von den Glykanketten der Muzine die Zugänglichkeit anderer Zucker der Glykanketten bewirkt. Diese könnten durch Galaktosidasen, *N*-Acetylhexosaminidasen, Fukosidasen und andere Enzyme der kommensalen Mikrobiota weiter degradiert werden. Weitere Substrate wären verfügbar und Muzin könnte schrittweise degradiert werden. Dies würde auch zur Zugänglichkeit des Epithels beitragen.
Im Gesamten hätte die Phagen-assoziierte Freisetzung von NanS-p Proteinen expandierende Auswirkungen für die Kolonisation pathogener *E. coli* in diesem Modell.

Um eine mögliche Beteiligung der NanS-p Proteine an der Infektion des menschlichen Dickdarms zu zeigen und einen tieferen Einblick in den Sialinsäurekatabolismus von pathogenen *E. coli* zu erhalten sollten idealerweise *in vivo* Untersuchungen erfolgen. Solche Untersuchungen in Tieren sind aufgrund mikrobieller, metabolischer, anatomischer oder biochemischer Unterschiede schwierig zu realisieren (Jubelin et al., 2018). Das Sialinsäuremuster ist anders, und humanpathogene Stämme sind in Tieren nicht infektiös. Es gibt bereits einige künstliche Darmmodelle, die Dynamik und Kompartimentierung berücksichtigen, jedoch sind Tiermodelle, die mit Streptomycin behandelt wurden, in der EHEC/STEC Forschung noch weit verbreitet (Jubelin et al., 2018). Möglich wären künstliche Systeme wie der „Simulator of the Human Intestinal Microbial Ecosystem (SHIME)“ Reaktor (Molly et al., 1993). Der SHIME Reaktor besteht aus fünf Bioreaktoren, die die Kompartimente Magen, Dünndarm, aufsteigender, transversaler und absteigender Dickdarm umfassen (Molly et al., 1993; Jubelin et al., 2018). Dieses System kann mit humanen Fäzes und Muzin-bedeckten Agarkugeln bestückt werden. Durch die Etablierung diverser *nanS*/*nanS*-p Deletionsmutanten in dieser Arbeit ist die Grundlage für solche Untersuchungen geschaffen. Wildtypstämme kombiniert mit ihren Mutantenstämmen, kommensalen *E. coli* und/oder Mitgliedern der humanen Darmmikrobiota könnten in entsprechenden Experimenten kombiniert eingesetzt werden.

Außerdem sollten Transkriptomanalysen und Untersuchungen zur Regulation von Virulenzfaktoren durch Sialinsäuren nach dem Vorbild der Arbeitsgruppe Le Bihan et al. (2015 & 2017) durchgeführt werden. Desweiteren stellt sich die Frage nach einer möglichen Funktion des freigesetzten Acetats. Um mögliche therapeutische Ansätze zu etablieren sollte analysiert werden wie NanS-p Proteine inhibiert werden können. Analysen zur Lokalisation von NanS-p Proteinen und der Beteiligung von Muzinasen, wie StcE, im Infektionsprozess stellen eine weitere Herausforderung für zukünftige Untersuchungen dar. Letztendlich stellt sich auch die Frage, in wie weit die neuen Erkenntnisse auch für andere *E. coli* Pathovare zu treffen.

Durch die Erkenntnisse dieser Arbeit entwickelten sich jedoch auch neue Fragestellungen zu anderen Bereichen. Bringen die NanS-p Proteine den Phagen einen Nutzen bei der Infektion von Bakterien? Sind die NanS-p Proteine am Phagen gebunden oder kommen sie im Phagenkapsid vor? In den Masterarbeiten „Investigation of the association of the phage-encoded protein 933Wp42 with the Stx2-phage 933W“ von Sara Farag (2016) und „Herstellung, Reinigung und Charakterisierung von infektiösen Lysaten des Shiga Toxin-kodierenden Bakteriophagen 933W“ von Katrin Haag (2016) am Fachgebiet Lebensmittelmikrobiologie und -hygiene konnte mit goldmarkierten Antikörpern gegen NanS-p1 keine Assoziation von NanS-p1 mit dem Stx2a-Phagen 933W detektiert werden. Auch wenn der Nachweis mit dieser Methodik nicht möglich war, sollten Untersuchungen zu dieser Thematik weiterhin berücksichtigt werden.

Interagieren NanS-p Proteine möglicherweise mit Stx? Es könnte gemutmaßt werden, dass NanS-p Proteine der Stx-Phagen eine Verstärkung der toxischen Wirkung von Stx bewirken. Der genaue Weg, wie Stx zu den Nieren und ins Gehirn gelangt, ist immer noch nicht komplett aufgeklärt (Schüller, 2011). NanS-p Proteine könnten in diesem Prozess eine Modifikation von Rezeptoren bewirken.

Desweiteren konnte in den NanS-p Aminosäuresequenzen eine Domäne (303) identifiziert werden, die für die Esterasefunktion verantwortlich ist, jedoch ist die Funktion einer zweiten (1737) unbekannt. Auch die Funktion der C-terminalen Region der NanS-p Proteine ist noch völlig unbekannt. Weitere enzymatische Aktivitä-

ten oder eine Funktion als Sialolektin könnten damit assoziiert sein. Solche Fusionsproteine sind aus der Virologie bekannt und bewirken eine spezifische Bindung an Sialinsäuren (Martin et al., 2003; Langereis et al., 2015).
Zuletzt ist auch an eine NanS-p Beteiligung bei der STEC Kolonisation von Wiederkäuern zu denken. Wiederkäuer sind das natürliche Habitat der STEC. Rinder sind die Tiere mit der diversesten Zusammensetzung an Sialinsäuren (Schauer, 2004, Lundblad, 2015). Möglicherweise haben die NanS-p bei der Kolonisation dieser Tiere eine selbe Funktion wie bei der Kolonisation des Menschen.

Zusammenfassend konnte in dieser Dissertation gezeigt werden, dass NanS-p Proteine von Prophagen kodiert werden, als Sialinsäureesterasen aktiv sind und von pathogenen *E. coli* zur Energiegewinnung genutzt werden. Das Auftreten multipler *nanS*-p Gene im Chromosom ist für einen Gen-Dosis Effekt nötig, der zu einer kurzen Generationszeit der bakteriellen Population bei der Vermehrung mit Neu5,9Ac_2 führt. Aufgrund der NanS-p Proteine können STEC *in vitro* mit Neu5,9Ac_2 wettbewerbsfähig sein und in Kombination mit einer Sialidase zur Bereitstellung von Sialinsäuren aus Muzin beitragen. Durch diese Ergebnisse und der Konstruktion diverser *nanS*/*nanS*-p Deletionsmutanten ist die Voraussetzung für zukünftige *in vivo* Untersuchungen, die zu einem tieferen Verständnis des Sialinsäurekatabolismus bei der EHEC Infektion des menschlichen Dickdarms beitragen, geschaffen.

Literaturverzeichnis

Die in Kapitel 2-4 verwendeten Quellen sind unmittelbar im entsprechenden Kapitel dargestellt. Die nachfolgenden Literaturquellen beziehen sich auf Kapitel 1 und 5.

Aiba, Y., Ishikawa, H., Shimizu, K., Noda, S., Kitada, Y., Sasaki, M., Koga, Y., 2002. Role of internalization in the pathogenicity of Shiga toxin-producing *Escherichia coli* infection in a gnotobiotic murine model. Microbial. Immunol. 46:723-731.

Aldick, T., Bielaszewska, M., Uhlin, B.E., Humpf, H.U., Wai, S.N., Karch, H., 2009. Vesicular stabilization and activity augmentation of enterohaemorrhagic *Escherichia coli* haemolysin. Mol. Microbial. 71:1496-1508.

Almagro-Moreno, S., Boyd, E.F., 2009. Insights into the evolution of sialic acid catabolism among bacteria. BMC Evol. Biol. 9:118.

Almagro-Moreno, S., Pruss, K., Taylor, R.K., 2015. Intestinal colonization dynamics of *Vibrio cholerae*. PLoS Pathog. 11:e1004787.

Amaral, M.M., Sacerdoti, F., Jancic, C., Repetto, H.A., Paton, A.W., Paton, J.C., Ibarra, C., 2013. Action of shiga toxin type-2 and subtilase cytotoxin on human microvascular endothelial cells. PLoS One 8:e70431.

Anonym a, 2012. Robert Koch-Institut: Infektionsepidemiologisches Jahrbuch für 2011. Abteilung für Infektionsepidemiologie, Robert Koch-Institut, Berlin, ISBN: 978-3-89606-246-8.

Anonym b, 2018. Basic Local Alignment Search Tool (Zugang 15. Dezember 2018). https://blast.ncbi.nlm.nih.gov/Blast.cgi.

Anonym c, 2018. National Center for Biotechnology Information (Zugang 14. Januar 2019). https://www.ncbi.nlm.nih.gov.

Arabyan, N., Park, D., Foutouhi, S., Weis, A.M., Huang, B.C., Williams, C.C., Desai, P., Shah, J., Jeannotte, R., Kong, N., Lebrilla, C.B., Weimer, B.C., 2016. *Salmonella* degrades the host glycocalyx leading to altered infection and glycan remodeling. Sci. Rep. 6:29525.

Arike, L., Hansson, G.C., 2016. The densely *O*-glycosylated MUC2 mucin protects the intestine and provides food for the commensal bacteria. J. Mol. Biol. 428:3221-3229.

Atuma, C., Strugala, V., Allen, A., Holm, L., 2001. The adherent gastrointestinal mucus gel layer: Thickness and physical state *in vivo*. Am. J. Physiol. Gastrointest. Liver Physiol. 280:922-929.

Baharoglu, Z., Mazel, D., 2014. SOS, the formidable strategy of bacteria against aggressions. FEMS Microbial. Rev. 38:1126-1145.

Bakshani, C.R., Morales-Garcia, A.L., Althaus, M., Wilcox, M.D., Pearson, J.P., Bythell, J.C., Burgess, J.G., 2018. Evolutionary conservation of the antimicrobial function of mucus: A first defence against infection. NPJ Biofilms Microbiomes 4:14.

Barondess, J.J., Beckwith, J., 1995. *bor* gene of phage lambda, involved in serum resistance, encodes a widely conserved outer membrane lipoprotein. J. Bacteriol. 177:1247-1253.

Battesti, A., Majdalani, N., Gottesman, S., 2011. The RpoS-mediated general stress response in *Escherichia coli*. Annu. Rev. Microbial. 65:189-213.

Bauwens, A., Bielaszewska, M., Kemper, B., Langehanenberg, P., von Bally, G., Reichelt, R., Mulac, D., Humpf, H.U., Friedrich, A.W., Kim, K.S., Karch, H., Müthing, J., 2011. Differential cytotoxic actions of Shiga toxin 1 and Shiga toxin 2 on microvascular and macrovascular endothelial cells. Thromb. Haemost. 105:515-528.

Bergstrom, K.S., Morampudi, V., Chan, J.M., Bhinder, G., Lau, J., Yang, H., Ma, C., Huang, T., Ryz, N., Sham, H.P., Zarepour, M., Zaph, C., Artis, D., Nair, M., Vallance, B.A., 2015. Goblet cell derived RELM-β recruits CD4+ T cells during infectious colitis to promote protective intestinal epithelial cell proliferation. PLoS Pathog. 11:e1005108.

Bertin, Y., Chaucheyras-Durand, F., Robbe-Masselot, C., Durand, A., de la Foye, A., Harel, J., Cohen, P.S., Conway, T., Forano, E., Martin, C., 2013. Carbohydrate utilization by enterohaemorrhagic *Escherichia coli* O157:H7 in bovine intestinal content. Environ. Microbial. 15:610-622.

Berry, J., Summer, E.J., Struck, D.K., Young, R., 2008. The final step in the phage infection cycle: The Rz and Rz1 lysis proteins link the inner and outer membranes. Mol. Microbiol. 70:341-351.

Betz, J., Dorn, I., Kouzel, I.U., Bauwens, A., Meisen, I., Kemper, B., Bielaszewska, M., Mormann, M., Weymann, L., Sibrowski, W., Karch, H., Schlenke, P., Müthing, J., 2016. Shiga toxin of enterohaemorrhagic *Escherichia coli* directly injures developing human erythrocytes. Cell Microbial. 18:1339-1348.

Bhowmick, R., Ghosal, A., Das, B., Koley, H., Saha, D.R., Ganguly, S., Nandy, R.K., Bhadra, R.K., Chatterjee, N.S., 2008. Intestinal adherence of *Vibrio cholerae* involves a coordinated interaction between colonization factor GbpA and mucin. Infect. Immun. 76:4968-4977.

Bielaszewska, M., Karch, H., 2005. Consequences of enterohaemorrhagic *Escherichia coli* infection for the vascular endothelium. Thromb. Haemost. 94:312-318.

Bielaszewska, M., Mellmann, A., Zhang, W., Köck, R., Fruth, A., Bauwens, A., Peters, G., Karch, H., 2011. Characterisation of the *Escherichia coli* strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: A microbiological study. Lancet Infect. Dis. 11:671-676.

Bielaszewska, M., Rüter, C., Bauwens, A., Greune, L., Jarosch, K.A., Steil, D., Zhang, W., He, X., Lloubes, R., Fruth, A., Kim, K.S., Schmidt, M.A., Dobrindt, U., Mellmann, A., Karch, H., 2017. Host cell interactions of outer membrane vesicle-associated virulence factors of enterohemorrhagic *Escherichia coli* O157: Intracellular delivery, trafficking and mechanisms of cell injury. PLoS Pathog. 13:e1006159.

Bielaszewska, M., Sinha, B., Kuczius, T., Karch, H., 2005. Cytolethal distending toxin from Shiga toxin-producing *Escherichia coli* O157 causes irreversible G2/M arrest, inhibition of proliferation, and death of human endothelial cells. Infect. Immun. 73:552-562.

Booth, B.A., Boesman-Finkelstein, M., Finkelstein, R.A., 1983. *Vibrio cholerae* soluble hemagglutinin/protease is a metalloenzyme. Infect. Immun. 42:639-644.

Brunder, W., Schmidt, H., Karch, H., 1996. KatP, a novel catalase-peroxidase encoded by the large plasmid of enterohaemorrhagic *Escherichia coli* O157:H7. Microbiology 142:3305-3315.

Brunder, W., Schmidt, H., Karch, H., 1997. EspP, a novel extracellular serine protease of enterohaemorrhagic *Escherichia coli* O157:H7 cleaves human coagulation factor V. Mol. Microbial. 24:767-778.

Brzuszkiewicz, E., Gottschalk, G., Ron, E., Hacker, J., Dobrindt, U., 2009. Adaptation of pathogenic *E. coli* to various niches: Genome flexibility is the key. Genome Dyn. 6:110-125.

Brzuszkiewicz, E., Thürmer, A., Schuldes, J., Leimbach, A., Liesegang, H., Meyer, F.D., Boelter, J., Petersen, H., Gottschalk, G., Daniel, R., 2011. 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. Microbial. 193:883-891.

Cahill, J., Rajaure, M., O'Leary, C., Sloan, J., Marrufo, A., Holt, A., Kulkarni, A., Hernandez, O., Young, R., 2017. Genetic analysis of the Lambda spanins Rz and Rz1: Identification of functional domains. G3 (Bethesda) 7:741-753.

Calderwood, S.B., Mekalanos, J.J., 1987. Iron regulation of Shiga-like toxin expression in *Escherichia coli* is mediated by the *fur* locus. J. Bacteriol. 169:4759-4764.

Cherepanov, P.P., Wackernagel, W., 1995. 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.

Chang, D.E., Smalley, D.J., Tucker, D.L., Leatham, M.P., Norris, W.E., Stevenson, S.J., Anderson, A.B., Grissom, J.E., Laux, D.C., Cohen, P.S., Conway, T., 2004. Carbon nutrition of *Escherichia coli* in the mouse intestine. Proc. Natl. Acad. Sci. U.S.A. 101:7427-7432.

Carlson-Banning, K.M., Sperandio, V., 2016. Catabolite and oxygen regulation of enterohemorrhagic *Escherichia coli* virulence. MBio 7:e01852-16.

Chapman, P.A., Cornell, J., Green, C., 2000. Infection with verocytotoxin-producing *Escherichia coli* O157 during a visit to an inner city open farm. Epidemiol. Infect. 125: 531-536.

Chong, Y., Fitzhenry, R., Heuschkel, R., Torrente, F., Frankel, G., Phillips, A.D., 2007. Human intestinal tissue tropism in *Escherichia coli* O157:H7 - initial colonization of terminal ileum and Peyer's patches and minimal colonic adhesion *ex vivo*. Microbiology 153:794-802.

Condemine, G., Berrier, C., Plumbridge, J., Ghazi, A., 2005. Function and expression of an *N*-acetylneuraminic acid-inducible outer membrane channel in *Escherichia coli*. J. Bacteriol. 187:1959-1965.

Conway, T., Cohen, P.S., 2015. Commensal and pathogenic *Escherichia coli* metabolism in the gut. Microbial. Spectr. 3:BP-0006-2014.

Corfield, A.P., 2015. Mucins: A biologically relevant glycan barrier in mucosal protection. Biochim Biophys. Acta 1850:236-252.

Corfield, A.P., 2018. The interaction of the gut microbiota with the mucus barrier in health and disease in human. Microorganisms 6:78.

Corfield, A.P., Wagner, S.A., Clamp, J.R., Kriaris, M.S., Hoskins, L.C., 1992. Mucin degradation in the human colon: production of sialidase, sialate *O*-acetylesterase, *N*-acetylneuraminate lyase, arylesterase, and glycosulfatase activities by strains of fecal bacteria. Infect. Immun. 60:3971-3978.

Costard, S., Espejo, L., Groenendaal, H., Zagmutt, F.J., 2017. Outbreak-related disease burden associated with consumption of unpasteurized cow’s milk and cheese, United States, 2009-2014. Emerg. Infect. Dis. 23:957-964.

Creuzburg, K., Middendorf, B., Mellmann, A., Martaler, T., Holz, C., Fruth, A., Karch, H., Schmidt, H., 2011. Evolutionary analysis and distribution of type III effector genes in pathogenic *Escherichia coli* from human, animal and food sources. Environ. Microbial. 13:439-452.

Crociani, F., Alessandrini, A., Mucci, M.M., Biavati, B., 1994. Degradation of complex carbohydrates by *Bifidobacterium* spp. Int. J. Food Microbial. 24:199-210.

Crost, E.H., Tailford, L.E., Le Gall, G., Fons, M., Henrissat, B., Juge, N., 2013. Utilisation of mucin glycans by the human gut symbiont *Ruminococcus gnavus* is strain-dependent. PLoS One 8:e76341.

Das, A., 1992. How the phage lambda *N* gene product suppresses transcription termination: Communication of RNA polymerase with regulatory proteins mediated by signals in nascent RNA. J. Bacteriol. 174:6711-6716.

Datsenko, K.A., Wanner, B.L., 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc. Natl. Acad. Sci. USA. 97:6640-6645.

Derrien, M., Vaughan, E.E., Plugge, C.M., de Vos, W.M., 2004. *Akkermansia muciniphila gen*. *nov*., *sp. nov.*, a human intestinal mucin-degrading bacterium. Int. J. Syst. Evol. Microbial. 54:1469-1476.

Dobrindt, U., 2005. (Patho-)Genomics of *Escherichia coli*. Int. J. Med. Microbial. 295:357-371.

Elliott, S.J., Sperandio, V., Girón, J.A., Shin, S., Mellies, J.L., Wainwright, L., Hutcheson, S.W., McDaniel, T.K., Kaper, J.B., 2000. The locus of enterocyte effacement (LEE)-encoded regulator controls expression of both LEE- and non-LEE-encoded virulence factors in enteropathogenic and enterohemorrhagic *Escherichia coli*. Infect. Immun. 68:6115-6126.

Endo, Y., Tsurugi, K., Yutsudo, T., Takeda, Y., Ogasawara, T., Igarashi, K., 1988. Site of action of a Vero toxin (VT2) from *Escherichia coli* O157:H7 and of Shiga toxin on eukaryotic ribosomes. RNA *N*-glycosidase activity of the toxins. Eur. J. Biochem. 171:45-50.

Engevik, M.A., Yacyshyn, M.B., Engevik, K.A., Wang, J., Darien, B., Hassett, D.J., Yacyshyn, B.R., Worrell, R.T., 2015. Human *Clostridium difficile* infection: Altered mucus production and composition. Am. J. Physiol. Gastrointest. Liver Physiol. 308:510-524.

Epple, H.J., Kreusel, K.M., Hanski, C., Schulzke, J.D., Riecken, E.O., Fromm, M., 1997. Differential stimulation of intestinal mucin secretion by cholera toxin and carbachol. Pflugers Arch. 433:638-647.

Erdem, A.L., Avelino, F., Xicohtencatl-Cortes, J., Girón, J.A., 2007. Host protein binding and adhesive properties of H6 and H7 flagella of attaching and effacing *Escherichia coli*. J. Bacteriol. 189:7426-7435.

Fabich, A.J., Jones, S.A., Chowdhury, F.Z., Cernosek, A., Anderson, A., Smalley, D., McHargue, J.W., Hightower, G.A., Smith, J.T., Autieri, S.M., Leatham, M.P., Lins, J.J., Allen, R.L., Laux, D.C., Cohen, P.S., Conway, T., 2008. Comparison of carbon nutrition for pathogenic and commensal *Escherichia coli* strains in the mouse intestine. Infect. Immun. 76:1143-1152.

Falcão, J.P., Sharp, F., Sperandio, V., 2004. Cell-to-cell signaling in intestinal pathogens. Curr. Issues Intest. Microbial. 5:9-17.

Fellner, L., Huptas, C., Simon, S., Mühlig, A., Scherer, S., Neuhaus, K., 2016. Draft genome sequences of three european laboratory derivatives from enterohemorrhagic *Escherichia coli* O157:H7 strain EDL933, including two plasmids. Genome Announc. 4:e01331-15.

Feuerbaum, S., Saile, N., Pohlentz, G., Müthing, J., Schmidt, H., 2018. De-*O*-acetylation of mucin-derived sialic acids by recombinant NanS-p esterases of *Escherichia coli* O157:H7 strain EDL933. Int. J. Med. Microbial. 308:1113-1120.

Fox, J.T., Drouillard, J.S., Shi, X., Nagaraja, T.G., 2009. Effects of mucin and its carbohydrate constituents on *Escherichia coli* O157 growth in batch culture fermentations with ruminal or fecal microbial inoculum. J. Anim. Sci. 87:1304-1313.

Frank, C., Werber, D., Cramer, J.P., Askar, M., Faber, M., an der Heiden, M., Bernard, H., Fruth, A., Prager, R., Spode, A., Wadl, M., Zoufaly, A., Jordan, S., Kemper, M.J., Follin, P., Müller, L., King, L.A., Rosner, B., Buchholz, U., Stark, K., Krause, G., 2011. Epidemic profile of Shiga-toxin-producing *Escherichia coli* O104:H4 outbreak in Germany. N. Engl. J. Med. 365:1771-1780.

Freedman, S.B., Xie, J., Neufeld, M.S., Hamilton, W.L., Hartling, L., Tarr, P.I.; Alberta Provincial Pediatric Enteric Infection Team (APPETITE), Nettel-Aguirre, A., Chuck, A., Lee, B., Johnson, D., Currie, G., Talbot, J., Jiang, J., Dickinson, J., Kellner, J., MacDonald, J., Svenson, L., Chui, L., Louie, M., Lavoie, M., Eltorki, M., Vanderkooi, O., Tellier, R., Ali, S., Drews, S., Graham, T., Pang, X.L., 2016. Shiga toxin-producing *Escherichia coli* infection, antibiotics, and risk of developing Hemolytic Uremic Syndrome: A meta-analysis. Clin. Infect. Dis. 62:1251-1258.

Freter, R., Brickner, H., Botney, M., Cleven, D., Aranki, A., 1983. Mechanisms that control bacterial populations in continuous-flow culture models of mouse large intestinal flora. Infect. Immun. 39:676-685.

Friedman, D.I., Court, D.L., 1995. Transcription antitermination: The lambda paradigm updated. Mol. Microbial. 18:191-200.

Fitzhenry, R.J., Stevens, M.P., Jenkins, C., Wallis, T.S., Heuschkel, R., Murch, S., Thomson, M., Frankel, G., Phillips, A.D., 2003. Human intestinal tissue tropism of intimin epsilon O103 *Escherichia coli*. FEMS Microbial. Lett. 218:311-316.

Furniss, R.C.D., Low, W.W., Mavridou, D.A.I., Dagley, L.F., Webb, A.I., Tate, E.W., Clements, A., 2018. Plasma membrane profiling during enterohemorrhagic *E. coli* infection reveals that the metalloprotease StcE cleaves CD55 from host epithelial surfaces. J. Biol. Chem. 293:17188-17199.

Gagnon, M., Zihler Berner, A., Chervet, N., Chassard, C., Lacroix, C., 2013. Comparison of the Caco-2, HT-29 and the mucus-secreting HT29-MTX intestinal cell models to investigate *Salmonella* adhesion and invasion. J. Microbial. Methods 94:274-279.

Gamage, S.D., Patton, A.K., Hanson, J.F., Weiss, A.A., 2004. Diversity and host range of Shiga toxin-encoding phage. Infect. Immun. 72:7131-7139.

Gaytán, M.O., Martínez-Santos, V.I., Soto, E., González-Pedrajo, B., 2016. Type three secretion system in attaching and effacing pathogens. Front. Cell. Infect. Microbial. 6:129.

Grant, J., Wendelboe, A.M., Wendel, A., Jepson, B., Torres, P., Smelser, C., Rolfs, R.T., 2006. Spinach-associated *Escherichia coli* O157:H7 outbreak, Utah and New Mexico, 2006. Emerg. Infect. Dis. 14:1633-1636.

Grif, K., Dierich, M.P., Karch, H., Allerberger, F., 1998. Strain-specific differences in the amount of Shiga toxin released from enterohemorrhagic *Escherichia coli* O157 following exposure to subinhibitory concentrations of antimicrobial agents. Eur. J. Clin. Microbial. Infect. Dis. 17:761-766.

Grys, T.E., Siegel, M.B., Lathem, W.W., Welch, R.A., 2005. The StcE protease contributes to intimate adherence of enterohemorrhagic *Escherichia coli* O157:H7 to host cells. Infect. Immun. 73:1295-1303.

Grys, T.E., Walters, L.L., Welch, R.A., 2006. Characterization of the StcE protease activity of *Escherichia coli* O157:H7. J. Bacteriol. 188:4646-4653.

Guccione, E., Leon-Kempis Mdel, R., Pearson, B.M., Hitchin, E., Mulholland, F., van Diemen, P.M., Stevens, M.P., Kelly, D.J., 2008. Amino acid-dependent growth of *Campylobacter jejuni*: Key roles for aspartase (AspA) under microaerobic and oxygen-limited conditions and identification of AspB (Cj0762), essential for growth on glutamate. Mol. Microbial. 69:77-93.

Haiko, J., Westerlund-Wikström, B., 2013. The role of the bacterial flagellum in adhesion and virulence. Biology 2:1242-1267.

Hansson, G.C., Johansson, M.E., 2010. The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. Gut Microbes 1:51-54.

Hendrix, R.W., Lawrence, J.G., Hatfull, G.F., Casjens, S., 2000. The origins and ongoing evolution of viruses. Trends Microbial. 8:504-508.

Hengge-Aronis, R., 1993. Survival of hunger and stress: The role of *rpoS* in early stationary phase gene regulation in *E. coli*. Cell 72:165-168.

Herold, S., Siebert, J., Huber, A., Schmidt, H., 2005. Global expression of prophage genes in *Escherichia coli* O157:H7 strain EDL933 in response to norfloxacin. Antimicrob. Agents Chemother. 49:931-944.

Heuvelink, A.E., van Heerwaarden, C., Zwartkruis-Nahuis, J.T., van Oosterom, R., Edink, K., van Duvnhoven, Y.T., de Boer, E., 2002. *Escherichia coli* O157 infection associated with a petting zoo. Epidemiol. Infect. 129:295-302.

Hews, C.L., Tran, S.L., Wegmann, U., Brett, B., Walsham, A.D., Kavanaugh, D., Ward, N.J., Juge, N., Schüller, S., 2017. The StcE metalloprotease of enterohaemorrhagic *Escherichia coli* reduces the inner mucus layer and promotes adherence to human colonic epithelium *ex vivo*. Cell Microbiol. 19:e12717.

Hofreuter, D., Novik, V., Galán, J.E., 2008. Metabolic diversity in *Campylobacter jejuni* enhances specific tissue colonization. Cell Host Microbe 4:425-433.

Honish, L., Punja, N., Nunn, S., Nelson, D., Hislop, N., Gosselin, G., Stashko, N., Dittrich, D., 2017. *Escherichia coli* O157:H7 infections associated with contaminated pork products - Alberta, Canada, July-October 2014. Can. Commun. Dis. Rep. 43:21-24.

Hopkins, A.P., Hawkhead, J.A., Thomas, G.H., 2013. Transport and catabolism of the sialic acids *N*-glycolylneuraminic acid and 3-keto-3-deoxy-D-glycero-D-galactononic acid by *Escherichia coli* K-12. FEMS Microbial. Lett. 347:14-22.

Hoskins, L.C., Agustines, M., McKee, W.B., Boulding, E.T., Kriaris, M., Niedermeyer, G., 1985. Mucin degradation in human colon ecosystems. Isolation and properties of fecal strains that degrade ABH blood group antigens and oligosaccharides from mucin glycoproteins. J. Clin. Invest. 75:944-953.

Hoyer, L.L., Hamilton, A.C., Steenbergen, S.M., Vimr, E.R., 1992. Cloning, sequencing and distribution of the *Salmonella typhimurium* LT2 sialidase gene, *nanH*, provides evidence for interspecies gene transfer. Mol. Microbial. 6:873-884.

Huang, A., Friesen, J., Brunton, J.L., 1987. Characterization of a bacteriophage that carries the genes for production of Shiga-like toxin 1 in *Escherichia coli*. J. Bacteriol. 169:4308-4312.

Imamovic, L., Ballesté, E., Martínez-Castillo, A., García-Aljaro, C., Muniesa, M., 2016. Heterogeneity in phage induction enables the survival of the lysogenic population. Environ. Microbial. 18:957-969.

In, J., Foulke-Abel, J., Zachos, N.C., Hansen, A.M., Kaper, J.B., Bernstein, H.D., Halushka, M., Blutt, S., Estes, M.K., Donowitz, M., Kovbasnjuk, O., 2016. Enterohemorrhagic *Escherichia coli* reduce mucus and intermicrovillar bridges in human stem cell-derived colonoids. Cell Mol. Gastroenterol. Hepatol. 2:48-62.

Iversen, H., Lindbäck, T., L'Abée-Lund, T.M., Roos, N., Aspholm, M., Stenfors Arnesen, L., 2015. The gut bacterium *Bacteroides thetaiotaomicron* influences the virulence potential of the enterohemorrhagic *Escherichia coli* O103:H25. PLoS One 10:e0118140.

Isogai, E., Isogai, H., Kimura, K., Hayashi, S., Kubota, T., Fujii, N., Takeshi, K., 1998. Role of tumor necrosis factor alpha in gnotobiotic mice infected with an *Escherichia coli* O157:H7 strain. Infect. Immun. 66:197-202.

Jacewicz, M., Clausen, H., Nudelman, E., Donohue-Rolfe, A., Keusch, G.T., 1986. Pathogenesis of *shigella* diarrhea. XI. Isolation of a *Shigella* toxin-binding glycolipid from rabbit jejunum and HeLa cells and its identification as globotriaosylceramide. J. Exp. Med. 163:1391-1404.

Jandhyala, D.M., Rogers, T.J., Kane, A., Paton, A.W., Paton, J.C., Thorpe, C.M., 2010. Shiga toxin 2 and flagellin from shiga-toxigenic *Escherichia coli* superinduce interleukin-8 through synergistic effects on host stress-activated protein kinase activation. Infect. Immun. 78:2984-2994.

Janka, A., Bielaszewska, M., Dobrindt, U., Greune, L., Schmidt, M.A., Karch, H., 2003. Cytolethal distending toxin gene cluster in enterohemorrhagic *Escherichia coli* O157:H^- and O157:H7: Characterization and evolutionary considerations. Infect. Immun. 71:3634-3638.

Jarvis, K.G., Girón, J.A., Jerse, A.E., McDaniel, T.K., Donnenberg, M.S., Kaper, J.B., 1995. Enteropathogenic *Escherichia coli* contains a putative type III secretion system necessary for the export of proteins involved in attaching and effacing lesion formation. Proc. Natl. Acad. Sci. U.S.A. 92:7996-8000.

Jerse, A.E., Yu, J., Tall, B.D., Kaper, J.B., 1990. A genetic locus of enteropathogenic *Escherichia coli* necessary for the production of attaching and effacing lesions on tissue culture cells. Proc. Natl. Acad. Sci. U.S.A. 87: 7839-7843.

Johannes, L., Römer, W., 2010. Shiga toxins - from cell biology to biomedical applications. Nat. Rev. Microbial. 8:105-116.

Johansson, M.E., Ambort, D., Pelaseyed, T., Schütte, A., Gustafsson, J.K., Ermund, A., Subramani, D.B., Holmén-Larsson, J.M., Thomsson, K.A., Bergström, J.H., van der Post, S., Rodriguez-Piñeiro, A.M., Sjövall, H., Bäckström, M., Hansson, G.C., 2011. Composition and functional role of the mucus layers in the intestine. Cell Mol. Life Sci. 68:3635-3641.

Johansson, M.E., Hansson, G.C., 2013. Mucus and the goblet cell. Dig. Dis. 31:305-309.

Johansson, M.E., Phillipson, M., Petersson, J., Velcich, A., Holm, L., Hansson, G.C., 2008. The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. Proc. Natl. Acad. Sci. U.S.A. 105:15064-15069.

Jonckheere, N., Skrypek, N., Frénois, F., Van Seuningen, I., 2013. Membrane-bound mucin modular domains: from structure to function. Biochimie 95:1077-1086.

Jordan, D.M., Cornick, N., Torres, A.G., Dean-Nystrom, E.A., Kaper, J.B., Moon, H.W., 2004. Long polar fimbriae contribute to colonization by *Escherichia coli* O157:H7 *in vivo*. Infect. Immun. 72:6168-6171.

Jores, J., Rumer, L., Wieler, L.H., 2004. Impact of the locus of enterocyte effacement pathogenicity island on the evolution of pathogenic *Escherichia coli*. Int. J. Med. Microbiol. 294:103-113.

Jubelin, G., Desvaux, M., Schüller, S., Etienne-Mesmin, L., Muniesa, M., Blanquet-Diot, S., 2018. Modulation of enterohaemorrhagic *Escherichia coli* survival and virulence in the human gastrointestinal tract. Microorganisms 6:E115.

Juge, N., 2012. Microbial adhesins to gastrointestinal mucus. Trends Microbial. 20:30-39.

Juge, N., Tailford, L., Owen, C.D., 2016. Sialidases from gut bacteria: A mini-review. Biochem. Soc. Trans. 44:166-175.

Kalivoda, K.A., Steenbergen, S.M., Vimr, E.R., 2013. Control of the *Escherichia coli* sialoregulon by transcriptional repressor NanR. J. Bacteriol. 195:4689-4701.

Kalivoda, K.A., Steenbergen, S.M., Vimr, E.R., Plumbridge, J., 2003. Regulation of sialic acid catabolism by the DNA binding protein NanR in *Escherichia coli*. J. Bacteriol. 185:4806-4815.

Karch, H., Schmidt, H., Janetzki-Mittmann, C., Scheef, J., Kröger, M., 1999. Shiga toxins even when different are encoded at identical positions in the genomes of related temperate bacteriophages. Mol. Gen Genet. 262:600-607.

Karch, H., Tarr, P.I., Bielaszewska, M., 2005. Enterohaemorrhagic *Escherichia coli* in human medicine. Int. J. Med. Microbial. 295:405-418.

Karpman, D., Connell, H., Svensson, M., Scheutz, F., Alm, P., Svanborg, C., 1997. The role of lipopolysaccharide and Shiga-like toxin in a mouse model of *Escherichia coli* O157:H7 infection. J. Infect. Dis. 175:611-620.

Karve, S.S., Pradhan, S., Ward, D.V., Weiss, A.A., 2017. Intestinal organoids model human responses to infection by commensal and Shiga toxin producing *Escherichia coli.* PLoS One 12:e0178966.

Kenny, B., DeVinney, R., Stein, M., Reinscheid, D.J., Frey, E.A., Finlay, B.B. 1997. Enteropathogenic *E. coli* (EPEC) transfers its receptor for intimate adherence into mammalian cells. Cell 91:511-520.

Kim, J.C., Yoon, J.W., Kim, C.H., Park, M.S., Cho, S.H., 2012. Repression of flagella motility in enterohemorrhagic *Escherichia coli* O157:H7 by mucin components. Biochem. Biophys. Res. Commun. 423:789-792.

King, L.A., Nogareda, F., Weill, F.X., Mariani-Kurkdjian, P., Loukiadis, E., Gault, G., Jourdan-DaSilva, N., Bingen, E., Macé, M., Thevenot, D., Ong, N., Castor, C., Noël, H., Van Cauteren, D., Charron, M., Vaillant, V., Aldabe, B., Goulet, V., Delmas, G., Couturier, E., Le Strat, Y., Combe, C., Delmas, Y., Terrier, F., Vendrely, B., Rolland, P., de Valk, H., 2012. Outbreak of Shiga toxin-producing *Escherichia coli* O104:H4 associated with organic fenugreek sprouts, France, June 2011. Clin. Infect. Dis. 54:1588-1594.

Köhler, B., Karch, H., Schmidt, H., 2000. Antibacterials that are used as growth promoters in animal husbandry can affect the release of Shiga-toxin-2-converting bacteriophages and Shiga toxin 2 from *Escherichia coli* strains. Microbiology 146:1085-1090.

Kolling, G.L., Matthews, K.R., 1999. Export of virulence genes and Shiga toxin by membrane vesicles of *Escherichia coli* O157:H7. Appl. Environ. Microbial. 65:1843-1848.

Krupovic, M., Bamford, D.H., 2008. Holin of bacteriophage lambda: Structural insights into a membrane lesion. Mol. Microbial. 69:781-783.

Langereis, M.A., Bakkers, M.J., Deng, L., Padler-Karavani. V., Vervoort. S.J., Hulswit, R.J., van Vliet, A.L., Gerwig, G.J., de Poot, S.A., Boot, W., van Ederen, A.M., Heesters, B.A., van der Loos, C.M., van Kuppeveld, F.J., Yu, H., Huizinga, E.G., Chen, X., Varki, A., Kamerling, J.P., de Groot, R.J., 2015. Complexity and diversity of the mammalian sialome revealed by *Nidovirus* virolectins. Cell Rep. 11:1966-1978.

Larsson, J.M., Karlsson, H., Sjövall, H., Hansson, G.C., 2009. A complex, but uniform *O*-glycosylation of the human MUC2 mucin from colonic biopsies analyzed by nanoLC/MSn. Glycobiology 19:756-766.

Lathem, W.W., Grys, T.E., Witowski, S.E., Torres, A.G., Kaper, J.B., Tarr, P.I., Welch, R.A., 2002. StcE, a metalloprotease secreted by *Escherichia coli* O157:H7, specifically cleaves C1 esterase inhibitor. Mol. Microbial. 45:277-288.

Latif, H., Li, H.J., Charusanti, P., Palsson, B.Ø., Aziz, R.K., 2014. A gapless, unambiguous genome sequence of the enterohemorrhagic *Escherichia coli* O157:H7 strain EDL933. Genome Announc. 14:2.

Le Bihan, G., Jubelin, G., Garneau, P., Bernalier-Donadille, A., Martin, C., Beaudry, F., Harel, J., 2015. Transcriptome analysis of *Escherichia coli* O157:H7 grown *in vitro* in the sterile-filtrated cecal content of human gut microbiota associated rats reveals an adaptive expression of metabolic and virulence genes. Microbes Infect. 17:23-33.

Le Bihan, G., Sicard, J.F., Garneau, P., Bernalier-Donadille, A., Gobert, A.P., Garrivier, A., Martin, C., Hay, A.G., Beaudry, F., Harel, J., Jubelin, G., 2017. The NAG sensor NagC regulates LEE gene expression and contributes to gut colonization by *Escherichia coli* O157:H7. Front. Cell Infect. Microbial. 7:134.

Lee, M.S., Kim, M.H., Tesh, V.L., 2013. Shiga toxins expressed by human pathogenic bacteria induce immune responses in host cells. J. Microbial. 51:724-730.

Legros, N., Dusny, S., Humpf, H.U., Pohlentz, G., Karch, H., Müthing, J., 2017. Shiga toxin glycosphingolipid receptors and their lipid membrane ensemble in primary human blood-brain barrier endothelial cells. Glycobiology 27:99-109.

Lencer, W.I., Reinhart, F.D., Neutra, M.R., 1990. Interaction of cholera toxin with cloned human goblet cells in monolayer culture. Am. J. Physiol. 258:96-102.

Levine, M.M., 1987. *Escherichia coli* that cause diarrhea: Enterotoxigenic, enteropathogenic, enteroinvasive, enterohemorrhagic, and enteroadherent. J. Infect. Dis. 155:377-389.

Lewis, S.B., Cook, V., Tighe, R., Schüller, S., 2015. Enterohemorrhagic *Escherichia coli* colonization of human colonic epithelium *in vitro* and *ex vivo*. Infect. Immun. 83:942-949.

Lewis, S.B., Prior, A., Ellis, S.J., Cook, V., Chan, S.S., Gelson, W., Schüller, S., 2016. Flagellin induces β-Defensin 2 in human colonic *ex vivo* infection with enterohemorrhagic *Escherichia coli*. Front. Cell Infect. Microbial. 6:68.

Li, Z., Bell, C., Buret, A., Robins-Browne, R., Stiel, D., O'Loughlin, E., 1993. The effect of enterohemorrhagic *Escherichia coli* O157:H7 on intestinal structure and solute transport in rabbits. Gastroenterology 104:467-474.

Lindberg, A.A., Brown, J.E., Strömberg, N., Westling-Ryd, M., Schultz, J.E., Karlsson, K.A., 1987. Identification of the carbohydrate receptor for Shiga toxin produced by *Shigella dysenteriae* type 1. J. Biol. Chem. 262:1779-1785.

Little, J.W., 1983. The SOS regulatory system: Control of its state by the level of RecA protease. J. Mol. Biol. 167:791-808.

Lundblad, A., 2015. Gunnar Blix and his discovery of sialic acids. Fascinating molecules in glycobiology. Ups. J. Med. Sci. 120:104-112.

Macfarlane, G.T., Gibson, G.R., 1991. Formation of glycoprotein degrading enzymes by *Bacteroides fragilis*. FEMS Microbial. Lett. 61:289-293.

Makino, K., Yokoyama, K., Kubota, Y., Yutsudo, C.H., Kimura, S., Kurokawa, K., Ishii, K., Hattori, M., Tatsuno, I., Abe, H., Iida, T., Yamamoto, K., Onishi, M., Hayashi, T., Yasunaga, T., Honda, T., Sasakawa, C., Shinagawa, H., 1999. Complete nucleotide sequence of the prophage VT2-Sakai carrying the verotoxin 2 genes of the enterohemorrhagic *Escherichia coli* O157:H7 derived from the Sakai outbreak. Genes Genet. Syst. 74:227-239.

Maltby, R., Leatham-Jensen, M.P., Gibson, T., Cohen, P.S., Conway, T., 2013. Nutritional basis for colonization resistance by human commensal *Escherichia coli* strains HS and Nissle 1917 against *E. coli* O157:H7 in the mouse intestine. PLoS One 8:e53957.

Mandal, C., Mandal, C., Chandra, S., Schauer, R., Mandal, C., 2012. Regulation of *O*-acetylation of sialic acids by sialate-*O*-acetyltransferase and sialate-*O*-acetylesterase activities in childhood acute lymphoblastic leukemia. Glycobiology 22:70-83.

Martin, L.T., Verhagen, A., Varki, A., 2003. Recombinant *Influenza C* hemagglutinin-esterase as a probe for sialic acid 9-*O*-acetylation. Methods Enzymol. 363:489-498.

Martinez, J., Steenbergen, S., Vimr, E., 1995. Derived structure of the putative sialic acid transporter from *Escherichia coli* predicts a novel sugar permease domain. J. Bacteriol. 177:6005-6010.

Matsushiro, A., Sato, K., Miyamoto, H., Yamamura, T., Honda, T., 1999. Induction of prophages of enterohemorrhagic *Escherichia coli* O157:H7 with norfloxacin. J. Bacteriol. 181:2257-2260.

McDaniel, T.K., Jarvis, K.G., Donnenberg, M.S., Kaper, J.B., 1995. A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. Proc. Natl. Acad. Sci. U.S.A. 92:1664-1668.

McDonald, N.D., Lubin, J.B., Chowdhury, N., Boyd, E.F., 2016. Host-derived sialic acids are an important nutrient source required for optimal bacterial fitness *in vivo* MBio 7:e02237-15.

Mellies, J.L., Elliott, S.J., Sperandio, V., Donnenberg, M.S., Kaper, J.B., 1999. The Per regulon of enteropathogenic *Escherichia coli*: Identification of a regulatory cascade and a novel transcriptional activator, the locus of enterocyte effacement (LEE)-encoded regulator (Ler). Mol. Microbial. 33:296-306.

Mellmann, A., Harmsen, D., Cummings, C.A., Zentz, E.B., Leopold, S.R., Rico, A., Prior, K., Szczepanowski, R., Ji, Y., Zhang, W., McLaughlin, S.F., Henkhaus, J.K., Leopold, B., Bielaszewska, M., Prager, R., Brzoska, P.M., Moore, R.L., Guenther, S., Rothberg, J.M., Karch, H., 2011. Prospective genomic characterization of the German enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid next generation sequencing technology. PLoS One 6:e22751.

Melton-Celsa, A.R., 2014. Shiga toxin (Stx) classification, structure, and function. Microbial. Spectr. 2:EHEC-0024-2013.

Melton-Celsa, A.R., O'Brien, A.D., 2014. New therapeutic developments against Shiga toxin-producing *Escherichia coli*. Microbial. Spectr. 2:EHEC-0013-2013.

Michino, H., Araki, K., Minami, S., Takaya, S., Sakai, N., Miyazaki, M., Ono, A., Yanagawa, H., 1999. Massive outbreak of *Escherichia coli* O157:H7 infection in schoolchildren in Sakai City, Japan, associated with consumption of white radish sprouts. Am. J. Epidemiol. 150:787-796.

Miranda, R.L., Conway, T., Leatham, M.P., Chang, D.E., Norris, W.E., Allen, J.H., Stevenson, S.J., Laux, D.C., Cohen, P.S., 2004. Glycolytic and gluconeogenic growth of *Escherichia coli* O157:H7 (EDL933) and *E. coli* K-12 (MG1655) in the mouse intestine. Infect. Immun. 72:1666-1676.

Molly, K., Vande Woestyne, M., Verstraete, W., 1993. Development of a 5-step multi-chamber reactor as a simulation of the human intestinal microbial ecosystem. Appl. Microbial. Biotechnol. 39:254-258.

Mondal, S.I., Islam, M.R., Sawaguchi, A., Asadulghani, M., Ooka, T., Gotoh, Y., Kasahara, Y., Ogura, Y., Hayashi, T., 2016. Genes essential for the morphogenesis of the Shiga toxin 2-transducing phage from *Escherichia coli* O157:H7. Sci. Rep. 6:39036.

Mora, A., Blanco, M., Blanco, J.E., Dahbi, G., López, C., Justel, P., Alonso, M.P., Echeita, A., Bernárdez, M.I., González, E.A., Blanco, J., 2007. Serotypes, virulence genes and intimin types of Shiga toxin (verocytotoxin)-producing *Escherichia coli* isolates from minced beef in Lugo (Spain) from 1995 through 2003. BMC Microbial. 7:13.

Moran, A.P., Gupta, A., Joshi, L., 2011. Sweet-talk: Role of host glycosylation in bacterial pathogenesis of the gastrointestinal tract. Gut 60:1412-1425.

Mühlenhoff, M., Eckhardt, M., Gerardy-Schahn, R., 1998. Polysialic acid: Three-dimensional structure, biosynthesis and function. Curr. Opin. Struct. Biol. 8:558-564.

Muraoka, W.T., Zhang, Q., 2011. Phenotypic and genotypic evidence for L-fucose utilization by *Campylobacter jejuni*. J. Bacteriol. 193:1065-1075.

Nassar, F.J., Rahal, E.A., Sabra, A., Matar, G.M., 2013. Effects of subinhibitory concentrations of antimicrobial agents on *Escherichia coli* O157:H7 Shiga toxin release and role of the SOS response. Foodborne Pathog. Dis. 10:805-812.

Nataro, J.P., Kaper J.B., 1998. Diarrheagenic *Escherichia coli*. Clin. Microbial. Rev. 11:142-201.

Newland, J.W., Strockbine, N.A., Miller, S.F., O'Brien, A.D., Holmes, R.K., 1985. Cloning of Shiga-like toxin structural genes from a toxin converting phage of *Escherichia coli*. Science 230:179-181.

Neely, M.N., Friedman, D.I., 1998a. Arrangement and functional identification of genes in the regulatory region of lambdoid phage H-19B, a carrier of a Shiga-like toxin. Gene 223:105-113.

Neely, M.N., Friedman, D.I., 1998b. Functional and genetic analysis of regulatory regions of coliphage H-19B: Location of shiga-like toxin and lysis genes suggest a role for phage functions in toxin release. Mol. Microbial. 28:1255-1267.

Njoroge, J.W., Nguyen, Y., Curtis, M.M., Moreira, C.G., Sperandio, V., 2012. Virulence meets metabolism: Cra and KdpE gene regulation in enterohemorrhagic *Escherichia coli.* MBio 3:e00280-12.

Nübling, S., Eisele, T., Stöber, H., Funk, J., Polzin, S., Fischer, L., Schmidt, H., 2014. Bacteriophage 933W encodes a functional esterase downstream of the Shiga toxin 2a operon. Int. J. Med. Microbial. 304:269-274.

O'Brien, A.D., Holmes, R.K., 1987. Shiga and Shiga-like toxins. Microbial. Rev. 51:206-220.

O'Brien, A.D., LaVeck, G.D., Thompson, M.R., Formal, S.B., 1982. Production of *Shigella dysenteriae* type 1-like cytotoxin by *Escherichia coli*. J. Infect. Dis. 146:763-769.

O'Brien, A.D., Melton, A.R., Schmitt, C.K., McKee, M.L., Batts, M.L., Griffin, D.E., 1993. Profile of *Escherichia coli* O157:H7 pathogen responsible for hamburger-borne outbreak of hemorrhagic colitis and hemolytic uremic syndrome in Washington. J. Clin. Microbial. 31:2799-2801.

O'Brien, A.D., Newland, J.W., Miller, S.F., Holmes, R.K., Smith, H.W., Formal, S.B., 1984. Shiga-like toxin-converting phages from *Escherichia coli* strains that cause hemorrhagic colitis or infantile diarrhea. Science 226:694-696.

Ohnishi, M., Terajima, J., Kurokawa, K., Nakayama, K., Murata, T., Tamura, K., Ogura, Y., Watanabe, H., Hayashi, T., 2002. Genomic diversity of enterohemorrhagic *Escherichia coli* O157 revealed by whole genome PCR scanning. Proc. Natl. Acad. Sci. U.S.A. 99:17043-17048.

Pacheco, A.R., Curtis, M.M., Ritchie, J.M., Munera, D., Waldor, M.K., Moreira, C.G., Sperandio, V., 2012. Fucose sensing regulates bacterial intestinal colonization. Nature 492:113-117.

Pacheco, A.R., Lazarus, J.E., Sit, B., Schmieder, S., Lencer, W.I., Blondel, C.J., Doench, J.G., Davis, B.M., Waldor, M.K., 2018. CRISPR screen reveals that EHEC's T3SS and Shiga toxin rely on shared host factors for infection. MBio 9: e01003-18.

Pacheco, A.R., Sperandio, V., 2015. Enteric pathogens exploit the microbiota-generated nutritional environment of the gut. Microbial. Spectr. 3:MBP-0001-2014.

Pang, T., Savva, C.G., Fleming, K.G., Struck, D.K., Young, R., 2009. Structure of the lethal phage pinhole. Proc. Natl. Acad. Sci. U.S.A. 106:18966-18971.

Papadopoulos, D., Schneider, D., Meier-Eiss, J., Arber, W., Lenski, R.E., Blot, M., 1999. Genomic evolution during a 10,000-generation experiment with bacteria. Proc. Natl. Acad. Sci. U.S.A. 96:3807-3812.

Paton, A.W., Paton, J.C., 2010. *Escherichia coli* Subtilase cytotoxin. Toxins 2:215-228.

Paton, A.W., Ratcliff, R.M., Doyle, R.M., Seymour-Murray, J., Davos, D., Lanser, J.A., Paton, J.C., 1996. Molecular microbiological investigation of an outbreak of hemolytic-uremic syndrome caused by dry fermented sausage contaminated with Shiga-like toxin-producing *Escherichia coli*. J. Clin. Microbial. 34:1622-1627.

Paton, A.W., Srimanote, P., Woodrow, M.C., Paton, J.C., 2001. Characterization of Saa, a novel autoagglutinating adhesin produced by locus of enterocyte effacement-negative Shiga-toxigenic *Escherichia coli* strains that are virulent for humans. Infect. Immun. 69:6999-7009.

Pearce, O.M., Läubli, H., 2016. Sialic acids in cancer biology and immunity. Glycobiology 26:111-128.

Pelaseyed, T., Bergström, J.H., Gustafsson, J.K., Ermund, A., Birchenough, G.M., Schütte, A., van der Post, S., Svensson, F., Rodríguez-Piñeiro, A.M., Nyström, E.E., Wising, C., Johansson, M.E., Hansson, G.C., 2014. The mucus and mucins of the goblet cells and enterocytes provide the first defense line of the gastrointestinal tract and interact with the immune system. Immunol. Rev. 260:8-20.

Pereira, F.C., Berry, D., 2017. Microbial nutrient niches in the gut. Environ. Microbial. 19:1366-1378.

Peri, K.G., Goldie, H., Waygood, E.B., 1990. Cloning and characterization of the *N*-acetylglucosamine operon of *Escherichia coli*. Biochem. Cell Biol. 68:123-137.

Perna, N.T., Mayhew, G.F., Pósfai, G., Elliott, S., Donnenberg, M.S., Kaper, J.B., Blattner, F.R., 1998. Molecular evolution of a pathogenicity island from enterohemorrhagic *Escherichia coli* O157:H7. Infect. Immun. 66:3810-3817.

Perna, N.T., Plunkett, G. 3rd, Burland, V., Mau, B., Glasner, J.D., Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A., Pósfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grotbeck, E.J., Davis, N.W., Lim, A., Dimalanta, E.T., Potamousis, K.D., Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C., Welch, R.A., Blattner, F.R., 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. Nature 409:529-533.

Phillips, A.D., Frankel, G., 2000. Intimin-mediated tissue specificity in enteropathogenic *Escherichia coli* interaction with human intestinal organ cultures. J. Infect. Dis. 181:1496-1500.

Phillips, A.D., Navabpour, S., Hicks, S., Dougan, G., Wallis, T., Frankel, G., 2000. Enterohaemorrhagic *Escherichia coli* O157:H7 target Peyer's patches in humans and cause attaching/effacing lesions in both human and bovine intestine. Gut 47:377-381.

Pifer, R., Sperandio, V., 2014. The interplay between the microbiota and enterohemorrhagic *Escherichia coli*. Microbial. Spectr. 2:EHEC-0015-2013.

Plunkett, G. 3rd, Rose, D.J., Durfee, T.J., Blattner, F.R., 1999. Sequence of Shiga toxin 2 phage 933W from *Escherichia coli* O157:H7: Shiga toxin as a phage late-gene product. J. Bacteriol. 181:1767-1778.

Plumbridge, J., Vimr, E., 1999. Convergent pathways for utilization of the amino sugars *N*-acetylglucosamine, *N*-acetylmannosamine, and *N*-acetylneuraminic acid by *Escherichia coli*. J. Bacteriol. 181:47-54.

Polzin, S., Huber, C., Eylert, E., Elsenhans, I., Eisenreich, W., Schmidt, H., 2013. Growth media simulating ileal and colonic environments affect the intracellular proteome and carbon fluxes of enterohemorrhagic *Escherichia coli* O157:H7 strain EDL933. Appl. Environ. Microbial. 79:3703-3715.

Ptashne, M., 1992. A genetic switch – Phage λ and higher organisms. Cell Press and Blackwell Scientific Publications, Cambridge, 2nd edition, ISBN 0865422095.

Rangel, A., Steenbergen, S.M., Vimr, E.R., 2016. Unexpected diversity of *Escherichia coli* sialate *O*-acetyl esterase NanS. J. Bacteriol. 198:2803-2809.

Rasko, D.A., Webster, D.R., Sahl, J.W., Bashir, A., Boisen, N., Scheutz, F., Paxinos, E.E., Sebra, R., Chin, C.S., Iliopoulos, D., Klammer, A., Peluso, P., Lee, L., Kislyuk, A.O., Bullard, J., Kasarskis, A., Wang, S., Eid, J., Rank, D., Redman, J.C., Steyert, S.R., Frimodt-Møller, J., Struve, C., Petersen, A.M., Krogfelt, K.A., Nataro, J.P., Schadt, E.E., Waldor, M.K., 2011. Origins of the *E. coli* strain causing an outbreak of hemolytic-uremic syndrome in Germany. N. Engl. J. Med. 365:709-717.

Rangarajan, E.S., Ruane, K.M., Proteau, A., Schrag, J.D., Valladares, R., Gonzalez, C.F., Gilbert, M., Yakunin, A.F., Cygler, M., 2011. Structural and enzymatic characterization of NanS (YjhS), a 9-*O*-Acetyl *N*-acetylneuraminic acid esterase from *Escherichia coli* O157:H7. Protein Sci. 20:1208-1219.

Recktenwald, J., Schmidt, H., 2002. The nucleotide sequence of Shiga toxin (Stx) 2e-encoding phage phiP27 is not related to other Stx phage genomes, but the modular genetic structure is conserved. Infect. Immun. 70:1896-1908.

Reddi, G., Pruss, K., Cottingham, K.L., Taylor, R.K., Almagro-Moreno, S., 2018. Catabolism of mucus components influences motility of *Vibrio cholerae* in the presence of environmental reservoirs. PLoS One 13:e0201383.

Reida, P., Wolff, M., Pöhls, H.W., Kuhlmann, W., Lehmacher, A., Aleksić, S., Karch, H., Bockemühl, J., 1994. An outbreak due to enterohaemorrhagic *Escherichia coli* O157:H7 in a children day care centre characterized by person-to-person transmission and environmental contamination. Zentralbl. Bakteriol. 281:534-543.

Riley, L.W., Remis, R.S., Helgerson, S.D., McGee, H.B., Wells, J.G., Davis, B.R., Hebert, R.J., Olcott, E.S., Johnson, L.M., Hargrett, N.T., Blake, P.A., Cohen, M.L., 1983. Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. N. Engl. J. Med. 308:681-685.

Robbe, C., Capon, C., Maes, E., Rousset, M., Zweibaum, A., Zanetta, J.P., Michalski, J.C., 2003. Evidence of regio-specific glycosylation in human intestinal mucins: Presence of an acidic gradient along the intestinal tract. J. Biol. Chem. 278:46337-46348.

Robinson, L.S., Lewis, W.G., Lewis, A.L., 2017. The sialate *O*-acetylesterase EstA from gut *Bacteroidetes* species enables sialidase-mediated cross-species foraging of 9-*O*-acetylated sialoglycans. J. Biol. Chem. 292:11861-11872.

Roca, A.I., Cox, M.M., 1997. RecA protein: Structure, function, and role in recombinational DNA repair. Prog. Nucleic Acid Res. Mol. Biol. 56:129-223.

Rogers, T.J., Paton, A.W., McColl, S.R., Paton, J.C., 2003. Enhanced CXC chemokine responses of human colonic epithelial cells to locus of enterocyte effacement-negative shiga-toxigenic *Escherichia coli*. Infect. Immun. 71:5623-5632.

Roxas, J.L., Koutsouris, A., Bellmeyer, A., Tesfay, S., Royan, S., Falzari, K., Harris, A., Cheng, H., Rhee, K.J., Hecht, G., 2010. Enterohemorrhagic *E. coli* alters murine intestinal epithelial tight junction protein expression and barrier function in a Shiga toxin independent manner. Lab. Invest. 90:1152-1168.

Saile, N., Schwarz, L., Eißenberger, K., Klumpp, J., Fricke, F.W., Schmidt, H., 2018. Growth advantage of *Escherichia coli* O104:H4 strains on 5-*N*-acetyl-9-*O*-acetyl neuraminic acid as a carbon source is dependent on heterogeneous phage-borne NanS-p esterases. Int. J. Med. Microbiol. 308:459-468.

Saile, N., Voigt, A., Kessler, S., Stressler, T., Klumpp, J., Fischer, L., Schmidt, H., 2016. *Escherichia coli* O157:H7 strain EDL933 harbors multiple functional prophage-associated genes necessary for the utilization of 5-*N*-acetyl-9-*O*-acetyl neuraminic acid as a growth substrate. Appl. Environ. Microbial. 82:5940-5950.

Samadder, P., Xicohtencatl-Cortes, J., Saldaña, Z., Jordan, D., Tarr, P.I., Kaper, J.B., Girón, J.A., 2009. The *Escherichia coli ycbQRST* operon encodes fimbriae with laminin-binding and epithelial cell adherence properties in Shiga-toxigenic *E. coli* O157:H7. Environ. Microbial. 11:1815-1826.

Samraj, A.N., Pearce, O.M., Läubli, H., Cittenden, A.N., Bergfeld, A.K., Banda, K., Gregg, C.J., Bingman, A.E., Secrest, P., Diaz, S.L., Varki, N.M., Varki, A., 2015. A red meat-derived glycan promotes inflammation and cancer progression. Proc. Natl. Acad. Sci. U.S.A. 112:542-547.

Samuel, J.E., Perera, L.P., Ward, S., O'Brien, A.D., Ginsburg, V., Krivan, H.C., 1990. Comparison of the glycolipid receptor specificities of Shiga-like toxin type II and Shiga-like toxin type II variants. Infect. Immun. 58:611-618.

Sandvig, K., Bergan, J., Dyve, A.B., Skotland, T., Torgersen, M.L., 2010. Endocytosis and retrograde transport of Shiga toxin. Toxicon 56:1181-1185.

Saxena, S.K., O'Brien, A.D., Ackerman, E.J., 1989. Shiga toxin, Shiga-like toxin II variant, and ricin are all single-site RNA *N*-glycosidases of 28S RNA when microinjected into *Xenopus oocytes*. J. Biol. Chem. 264:596-601.

Scavia, G., Gianviti, A., Labriola, V., Chiani, P., Maugliani, A., Michelacci, V., Minelli, F., Tozzoli, R., Caprioli, A., Morabito, S., 2018. A case of haemolytic uraemic syndrome (HUS) revealed an outbreak of Shiga toxin-2-producing *Escherichia coli* O26:H11 infection in a nursery, with long-lasting shedders and person-to-person transmission, Italy 2015. J. Med. Microbial. doi: 10.1099/jmm.0.000738.

Schauer, R., 2004. Sialic acids: Fascinating sugars in higher animals and man. Zoology (Jena) 107:49-64.

Schmidt, H., 2001. Shiga-toxin-converting bacteriophages. Res. Microbial. 152:687-695.

Schmidt, H., Henkel, B., Karch, H., 1997. A gene cluster closely related to type II secretion pathway operons of gram-negative bacteria is located on the large plasmid of enterohemorrhagic *Escherichia coli* O157 strains. FEMS Microbial. Lett. 148:265-272.

Schmidt, H., Karch, H., Beutin, L., 1994. The large-sized plasmids of enterohemorrhagic *Escherichia coli* O157 strains encode hemolysins which are presumably members of the *E. coli* alpha-hemolysin family. FEMS Microbiol. Lett. 117:189-196.

Schmidt, H., Scheef, J., Janetzki-Mittmann, C., Datz, M., Karch, H., 1997. An *ileX* tRNA gene is located close to the Shiga toxin II operon in enterohemorrhagic *Escherichia coli* O157 and non-O157 strains. FEMS Microbiol. Lett. 149:39-44.

Schnaar, R.L., Gerardy-Schahn, R., Hildebrandt, H., 2014. Sialic acids in the brain: Gangliosides and polysialic acid in nervous system development, stability, disease, and regeneration. Physiol. Rev. 94:461-518.

Schüller, S., 2011. Shiga toxin interaction with human intestinal epithelium. Toxins (Basel) 3:626-639.

Schüller, S., Phillips, A.D., 2010. Microaerobic conditions enhance type III secretion and adherence of enterohaemorrhagic *Escherichia coli* to polarized human intestinal epithelial cells. Environ. Microbial. 12:2426-2435.

Scotland, S.M., Smith, H.R., Willshaw, G.A., Rowe, B., 1983. Vero cytotoxin production in strain of *Escherichia coli* is determined by genes carried on bacteriophage. Lancet 2:216.

Segura, A., Bertoni, M., Auffret, P., Klopp, C., Bouchez, O., Genthon, C., Durand, A., Bertin, Y., Forano, E., 2018. Transcriptomic analysis reveals specific metabolic pathways of enterohemorrhagic *Escherichia coli* O157:H7 in bovine digestive contents. BMC Genomics 19:766.

Severi, E., Müller, A., Potts, J.R., Leech, A., Williamson, D., Wilson, K.S., Thomas, G.H., 2008. Sialic acid mutarotation is catalyzed by the *Escherichia coli* beta-propeller protein YjhT. J. Biol. Chem. 283:4841-4849.

Shen, Y., Kohla, G., Lrhorfi, A.L., Sipos, B., Kalthoff, H., Gerwig, G.J., Kamerling, J.P., Schauer, R., Tiralongo, J., 2004. *O*-acetylation and de-*O*-acetylation of sialic acids in human colorectal carcinoma. Eur. J. Biochem. 271:281-290.

Shi, W.X., Chammas, R., Varki, A., 1996. Regulation of sialic acid 9-*O*-acetylation during the growth and differentiation of murine erythroleukemia cells. J. Biol. Chem. 271:31517-31525.

Shigeno, T., Akamatsu, T., Fujimori, K., Nakatsuji, Y., Nagata, A., 2002. The clinical significance of colonoscopy in hemorrhagic colitis due to enterohemorrhagic *Escherichia coli* O157:H7 infection. Endoscopy 34:311-314.

Sicard, J.F., Le Bihan, G., Vogeleer, P., Jacques, M., Harel, J., 2017. Interactions of intestinal bacteria with components of the intestinal mucus. Front. Cell. Infect. Microbial. 7:387.

Silva, A.J., Pham, K., Benitez, J.A., 2003. Haemagglutinin/protease expression and mucin gel penetration in El Tor biotype *Vibrio cholerae*. Microbiology 149:1883-1891.

Smith, J., 2001. The social evolution of bacterial pathogenesis. Proc. Biol. Sci. 268:61-69.

Smith, H.W., Green, P., Parsell, Z., 1983. Vero cell toxins in *Escherichia coli* and related bacteria: Transfer by phage and conjugation and toxic action in laboratory animals, chickens and pigs. J. Gen. Microbial. 129:3121-3137.

Smith, D.L., Rooks, D.J., Fogg, P.C., Darby, A.C., Thomson, N.R., McCarthy, A.J., Allison, H.E., 2012. Comparative genomics of Shiga toxin encoding bacteriophages. BMC Genomics 13:311.

Snider, T.A., Fabich, A.J., Conway, T., Clinkenbeard, K.D., 2009. *E. coli* O157:H7 catabolism of intestinal mucin-derived carbohydrates and colonization. Vet. Microbial. 136:150-154.

Sonnenburg, E.D., Zheng, H., Joglekar, P., Higginbottom, S.K., Firbank, S.J., Bolam, D.N., Sonnenburg, J.L., 2010. Specificity of polysaccharide use in intestinal *Bacteroides* species determines diet-induced microbiota alterations. Cell 141:1241-1252.

Souza, J.M., Plumbridge, J.A., Calcagno, M.L., 1997. *N*-acetylglucosamine-6-phosphate deacetylase from *Escherichia coli*: Purification and molecular and kinetic characterization. Arch. Biochem. Biophys. 340:338-346.

Sperandio, V., Torres, A.G., Girón, J.A., Kaper, J.B., 2001. Quorum sensing is a global regulatory mechanism in enterohemorrhagic *Escherichia coli* O157:H7. J. Bacteriol. 183:5187-5197.

Sperandio, V., Torres, A.G., Jarvis, B., Nataro, J.P., Kaper, J.B., 2003. Bacteria-host communication: The language of hormones. Proc. Natl. Acad. Sci. U.S.A. 100:8951-8956.

Springer, S.A., Diaz, S.L., Gagneux, P., 2014. Parallel evolution of a self-signal: Humans and new world monkeys independently lost the cell surface sugar Neu5Gc. Immunogenetics 66:671-674.

Stahl, M., Friis, L.M., Nothaft, H., Liu, X., Li, J., Szymanski, C.M., Stintzi, A., 2011. L-fucose utilization provides *Campylobacter jejuni* with a competitive advantage. Proc. Natl. Acad. Sci. U.S.A. 108:7194-7199.

Stahl, M., Frirdich, E., Vermeulen, J., Badayeva, Y., Li, X., Vallance, B.A., Gaynor, E.C., 2016. The helical shape of *Campylobacter jejuni* promotes *in vivo* pathogenesis by aiding transit through intestinal mucus and colonization of crypts. Infect. Immun. 84:3399-3407.

Steenbergen, S.M., Jirik, J.L., Vimr, E.R., 2009. YjhS (NanS) is required for *Escherichia coli* to grow on 9-*O*-acetylated *N*-acetylneuraminic acid. J. Bacteriol. 191:7134-7139.

Stevens, M.P., Frankel, G.M., 2014. The locus of enterocyte effacement and associated virulence factors of enterohemorrhagic *Escherichia coli*. Microbial. Spectr. 2:EHEC-0007-2013.

Strockbine, N.A., Marques, L.R., Newland, J.W., Smith, H.W., Holmes, R.K., O'Brien, A.D., 1986. Two toxin-converting phages from *Escherichia coli* O157:H7 strain 933 encode antigenically distinct toxins with similar biologic activities. Infect. Immun. 53:135-140.

Szabady, R.L., Yanta, J.H., Halladin, D.K., Schofield, M.J., Welch, R.A., 2011. TagA is a secreted protease of *Vibrio cholerae* that specifically cleaves mucin glycoproteins. Microbiology 157:516-525.

Tailford, L.E., Crost, E.H., Kavanaugh, D., Juge, N., 2015. Mucin glycan foraging in the human gut microbiome. Front. Genet. 6:81.

Tarr, P.I., 2009. Shiga toxin-associated hemolytic uremic syndrome and thrombotic thrombocytopenic purpura: distinct mechanisms of pathogenesis. Kidney Int. Suppl. 112:29-32.

Tarr, P.I., Bilge, S.S., Vary, J.C. Jr, Jelacic, S., Habeeb, R.L., Ward, T.R., Baylor, M.R., Besser, T.E., 2000. Iha: A novel *Escherichia coli* O157:H7 adherence-conferring molecule encoded on a recently acquired chromosomal island of conserved structure. Infect. Immun. 68:1400-1407.

Tasteyre, A., Barc, M.C., Collignon, A., Boureau, H., Karjalainen, T., 2001. Role of FliC and FliD flagellar proteins of *Clostridium difficile* in adherence and gut colonization. Infect. Immun. 69:7937-7940.

Tatsuno, I., Horie, M., Abe, H., Miki, T., Makino, K., Shinagawa, H., Taguchi, H., Kamiya, S., Hayashi, T., Sasakawa, C., 2001. *toxB* gene on pO157 of enterohemorrhagic *Escherichia coli* O157:H7 is required for full epithelial cell adherence phenotype. Infect. Immun. 69:6660-6669.

Tesh, V.L., 2010. Induction of apoptosis by Shiga toxins. Future Microbial. 5:431-453.

te Loo, D.M., Monnens, L.A., van Der Velden, T.J., Vermeer, M.A., Preyers, F., Demacker, P.N., van Den Heuvel, L.P., van Hinsbergh, V.W., 2000. Binding and transfer of verocytotoxin by polymorphonuclear leukocytes in hemolytic uremic syndrome. Blood 95:3396-3402.

Tietze, E., Dabrowski, P.W., Prager, R., Radonic, A., Fruth, A., Auraß, P., Nitsche, A., Mielke, M., Flieger, A., 2015. Comparative genomic analysis of two novel sporadic Shiga toxin-producing *Escherichia coli* O104:H4 strains isolated 2011 in Germany. PLoS One 10:e0122074.

Toth, I., Cohen, M.L., Rumschlag, H.S., Riley, L.W., White, E.H., Carr, J.H., Bond, W.W., Wachsmuth, I.K., 1990. Influence of a 60-megadalton plasmid on adherence of *Escherichia coli* O157:H7 and genetic derivatives. Infect. Immun. 58:1223-1231.

Tse, C.M., In, J.G., Yin, J., Donowitz, M., Doucet, M., Foulke-Abel, J., Ruiz-Perez, F., Nataro, J.P., Zachos, N.C., Kaper, J.B., Kovbasnjuk, O., 2018. Enterohemorrhagic *E. coli* (EHEC)-secreted serine protease EspP stimulates electrogenic ion transport in human colonoid monolayers. Toxins (Basel) 10:351.

Tu, Q.V., McGuckin, M.A., Mendz, G.L., 2008. *Campylobacter jejuni* response to human mucin MUC2: modulation of colonization and pathogenicity determinants. J. Med. Microbial. 57:795-802.

Tu, X., Nisan, I., Yona, C., Hanski, E., Rosenshine, I., 2003. EspH, a new cytoskeleton-modulating effector of enterohaemorrhagic and enteropathogenic *Escherichia coli*. Mol. Microbial. 47:595-606.

Turroni, F., Özcan, E., Milani, C., Mancabelli, L., Viappiani, A., van Sinderen, D., Sela, D.A., Ventura, M., 2015. Glycan cross-feeding activities between bifidobacteria under *in vitro* conditions. Front. Microbial. 6:1030.

Tuttle, J., Gomez, T., Doyle, M.P., Wells, J.G., Zhao, T., Tauxe, R.V., Griffin, P.M., 1999. Lessons from a large outbreak of *Escherichia coli* O157:H7 infections: Insights into the infectious dose and method of widespread contamination of hamburger patties. Epidemiol. Infect. 122:185-192.

Tyler, J.S., Mills, M.J., Friedman, D.I., 2004. The operator and early promoter region of the Shiga toxin type 2-encoding bacteriophage 933W and control of toxin expression. J. Bacteriol. 186:7670-7679.

Unkmeir, A., Schmidt, H., 2000. Structural analysis of phage-borne *stx* genes and their flanking sequences in shiga toxin-producing *Escherichia coli* and *Shigella dysenteriae* type 1 strains. Infect. Immun. 68:4856-4864.

Varki, A., 2009. Multiple changes in sialic acid biology during human evolution. Glycoconj. J. 26:231-245.

Varki, A., Gagneux, P., 2012. Multifarious roles of sialic acids in immunity. Ann. N. Y. Acad. Sci. 1253:16-36.

Varki, A., Schnaar, R.L., Schauer, R., 2017. Sialic acids and other nonulosonic acids. Essentials of Glycobiology, 3rd edition. Cold Spring Harbor Laboratory Press. Bookshelf ID: NBK453082.

Vica Pacheco, S., García González, O., Paniagua Contreras, G.L., 1997. The *lom* gene of bacteriophage lambda is involved in *Escherichia coli* K12 adhesion to human buccal epithelial cells. FEMS Microbial. Lett. 156:129-132.

Vimal, D.B., Khullar, M., Gupta, S., Ganguly, N.K., 2000. Intestinal mucins: The binding sites for *Salmonella* Typhimurium. Mol. Cell Biochem. 204:107-117.

Vimr, E.R., 2013. Unified theory of bacterial sialometabolism: How and why bacteria metabolize host sialic acids. ISRN Microbial. 2013:816713.

Vimr, E.R., Kalivoda, K.A., Deszo, E.L., Steenbergen, S.M., 2004. Diversity of microbial sialic acid metabolism. Microbial. Mol. Biol. Rev. 68:132-153.

Vimr, E.R., Troy, F.A., 1985. Identification of an inducible catabolic system for sialic acids (nan) in *Escherichia coli*. J. Bacteriol. 164:845-853.

Vogt, S.L., Peña-Díaz, J., Finlay, B.B., 2015. Chemical communication in the gut: Effects of microbiota-generated metabolites on gastrointestinal bacterial pathogens. Anaerobe 34:106-115.

Wasik, B.R., Barnard, K.N., Ossiboff, R.J., Khedri, Z., Feng, K.H., Yu, H., Chen, X., Perez, D.R., Varki, A., Parrish, C.R., 2017. Distribution of *O*-acetylated sialic acids among target host tissues for Influenza virus. mSphere. 2:e00379-16.

Wong, C.S., Jelacic, S., Habeeb, R.L., Watkins, S.L., Tarr, P.I., 2000. The risk of the hemolytic-uremic syndrome after antibiotic treatment of *Escherichia coli* O157:H7 infections. N. Engl. J. Med. 342:1930-1936.

Xiaoli, L., Figler, H.M., Goswami Banerjee, K., Hayes, C.S., Dudley, E.G., 2018. Non-pathogenic *Escherichia coli* enhance Stx2a production of *E. coli* O157:H7 through both *bamA*-dependent and independent mechanisms. Front. Microbial. 9:1325.

Xicohtencatl-Cortes, J., Monteiro-Neto, V., Ledesma, M.A., Jordan, D.M., Francetic, O., Kaper, J.B., Puente, J.L., Girón, J.A., 2008. Intestinal adherence associated with

type IV pili of enterohemorrhagic *Escherichia coli* O157:H7. J. Clin. Invest. 117:3519-3529.

Xu, J., Bjursell, M.K., Himrod, J., Deng, S., Carmichael, L.K., Chiang, H.C., Hooper, L.V., Gordon, J.I., 2003. A genomic view of the human - *Bacteroides thetaiotaomicron* symbiosis. Science 299:2074-2076.

Xue, Y., Zhang, H., Wang, H., Hu, J., Du, M., Zhu, M.J., 2014. Host inflammatory response inhibits *Escherichia coli* O157:H7 adhesion to gut epithelium through augmentation of mucin expression. Infect. Immun. 82:1921-1930.

Yarnell, W.S., Roberts, J.W., 1992. The phage lambda gene *Q* transcription antiterminator binds DNA in the late gene promoter as it modifies RNA polymerase. Cell 69:1181-1189.

Yokoyama, K., Makino, K., Kubota, Y., Watanabe, M., Kimura, S., Yutsudo, C.H., Kurokawa, K., Ishii, K., Hattori, M., Tatsuno, I., Abe, H., Yoh, M., Iida, T., Ohnishi, M., Hayashi, T., Yasunaga, T., Honda, T., Sasakawa, C., Shinagawa, H., 2000. Complete nucleotide sequence of the prophage VT1-Sakai carrying the Shiga toxin 1 genes of the enterohemorrhagic *Escherichia coli* O157:H7 strain derived from the Sakai outbreak. Gene 258:127-139.

Yoon, J.W., Lim, J.Y., Park, Y.H., Hovde, C.J., 2005. Involvement of the *Escherichia coli* O157:H7 (pO157) *ecf* operon and lipid A myristoyl transferase activity in bacterial survival in the bovine gastrointestinal tract and bacterial persistence in farm water troughs. Infect. Immun. 73:2367-2378.

Young, K.T., Davis, L.M., Dirita, V.J., 2007. *Campylobacter jejuni*: Molecular biology and pathogenesis. Nat. Rev. Microbial. 5:665-679.

Zangari, T., Melton-Celsa, A.R., Panda, A., Boisen, N., Smith, M.A., Tatarov, I., De Tolla, L.J., Nataro, J.P., O'Brien, A.D., 2013. Virulence of the Shiga toxin type 2-expressing *Escherichia coli* O104:H4 German outbreak isolate in two animal models. Infect. Immun. 81:1562-1574.

Zhang, X., McDaniel, A.D., Wolf, L.E., Keusch, G.T., Waldor, M.K., Acheson, D.W., 2000. Quinolone antibiotics induce Shiga toxin-encoding bacteriophages, toxin production, and death in mice. J. Infect. Dis. 181:664-670.

Zweifel, C., Giezendanner, N., Corti, S., Krause, G., Beutin, L., Danuser, J., Stephan, R., 2010. Characteristics of Shiga toxin-producing *Escherichia coli* isolated from Swiss raw milk cheese within a 3-year monitoring program. J. Food Prot. 73:88-91.