AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Bacteriology

Unexpected Diversity of *Escherichia coli* Sialate *O*-Acetyl Esterase NanS

Ariel Rangel,* Susan M. Steenbergen, Eric R. Vimr

Laboratory of Sialobiology, Department of Pathobiology, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

ABSTRACT

The sialic acids (*N*-acylneuraminates) are a group of nine-carbon keto-sugars existing mainly as terminal residues on animal glycoprotein and glycolipid carbohydrate chains. Bacterial commensals and pathogens exploit host sialic acids for nutrition, adhesion, or antirecognition, where *N*-acetyl- or *N*-glycolylneuraminic acids are the two predominant chemical forms of sialic acids. Each form may be modified by acetyl esters at carbon position 4, 7, 8, or 9 and by a variety of less-common modifications. Modified sialic acids produce challenges for colonizing bacteria, because the chemical alterations to *N*-acetylneuraminic acid (Neu5Ac) confer increased resistance to sialidase and aldolase activities essential for the catabolism of host sialic acids. Bacteria with *O*-acetyl sialate esterase(s) utilize acetylated sialic acids for growth, thereby gaining a presumed metabolic advantage over competitors lacking this activity. Here, we demonstrate the esterase activity of *Escherichia coli* NanS after purifying it as a C-terminal HaloTag fusion. Using a similar approach, we show that *E. coli* strain O157:H7 Stx prophage or prophage remnants invariably include paralogs of *nanS* often located downstream of the Shiga-like toxin genes. These paralogs may include sequences encoding N- or C-terminal domains of unknown function where the NanS domains can act as sialate *O*-acetyl esterases, as shown by complementation of an *E. coli* strain K-12 *nanS* mutant and the unimpaired growth of an *E. coli* O157 *nanS* mutant on *O*-acetylated sialic acid. We further demonstrate that *nanS* homologs in *Streptococcus* spp. also encode active esterase, demonstrating an unexpected diversity of bacterial sialate *O*-acetyl esterase.

IMPORTANCE

The sialic acids are a family of over 40 naturally occurring 9-carbon keto-sugars that function in a variety of host-bacterium interactions. These sugars occur primarily as terminal carbohydrate residues on host glycoproteins and glycolipids. Available evidence indicates that diverse bacterial species use host sialic acids for adhesion or as sources of carbon and nitrogen. Our results show that the catabolism of the diacetylated form of host sialic acid requires a specialized esterase, NanS. Our results further show that *nanS* homologs exist in bacteria other than *Escherichia coli*, as well as part of toxigenic *E. coli* prophage. The unexpected diversity of these enzymes suggests new avenues for investigating host-bacterium interactions. Therefore, these original results extend our previous studies of *nanS* to include mucosal pathogens, prophage, and prophage remnants. This expansion of the *nanS* superfamily suggests important, although as-yet-unknown, functions in host-microbe interactions.

Our laboratory discovered (1, 2) and later characterized the transport and metabolism of exogenous *N*-acetylneuraminic acid (Neu5Ac) in *Escherichia coli*, *Haemophilus influenzae*, and *Pasteurella multocida* (3–6), providing also the first evidence that these processes are essential for animal disease. The essential role of sialic acid transport in disease has been independently confirmed (7). Other studies, using a variety of bacterial species, support the idea that the metabolism of host-derived sialic acid is important to colonization or disease (see reference 8 and references cited therein). Although these studies have not been independently confirmed, the combined evidence suggests that the metabolism of host-derived sialic acids plays an at least minor to potentially defining role in different bacterial infections. Therefore, blocking bacterial access to host sialic acids could have therapeutic benefit (9, 10). However, all studies to date have been focused on the metabolism of Neu5Ac, the most common sialic acid, despite evidence that it is just one of several chemically distinct forms of sialic acid attached to host glycoconjugates (8). A recent study from this laboratory showed that *Escherichia coli* uses *O*-acetylated sialic acid as a sole source of carbon (11), and that this metabolic capability is dependent on a sialate, *O*-acetyl esterase (NanS), which is coordinately regulated by the NanR transcription repressor controlling expression of the sialoregulon (5, 12). The function of NanS was suggested by a low primary structural identity to *Rhodopirellula baltica* Sh1 AxeA, an acetyl xylan esterase (11).

The sialoregulon in *E. coli* is a group of 10 genes organized in three operons whose expression is controlled by the GntR superfamily regulatory protein NanR (12). The current study presents a biochemical characterization of *E. coli* strain K-12 NanS and identifies functional paralogs of this esterase in *E. coli* that also express Shiga or Shiga-like toxins. We further show that *Streptococcus* spp. express *nanS* with resulting *O*-acetylesterase activity. The combined results support the hypothesis that sialometabolism of sialic

Received 29 February 2016 **Accepted** 23 July 2016

Accepted manuscript posted online 1 August 2016

Citation Rangel A, Steenbergen SM, Vimr ER. 2016. Unexpected diversity of *Escherichia coli* sialate *O*-acetyl esterase NanS. J Bacteriol 198:2803–2809. doi:10.1128/JB.00189-16.

Editor: V. J. DiRita, Michigan State University

Address correspondence to Eric R. Vimr, ervimr@illinois.edu.

* Present address: Ariel Rangel, Thermo Fisher Scientific, Rockford, Illinois, USA.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/JB.00189-16.

acids other than Neu5Ac might be important to a wide range of host-bacterium interactions (8).

MATERIALS AND METHODS

Bacterial strains and growth conditions. *E. coli* K-12 and its Δ*nanS* mutant derivative have been described (11). *E. coli* DH5α is a laboratory stock strain. *E. coli* K-12 strain KRX was purchased from Promega. This strain contains an insertion of the gene for T7 polymerase in the chromosomal rhamnose operon. Thus, the polymerase is induced by the introduction of rhamnose to a growing culture. Bacterial growth was at 37°C with vigorous aeration in a rotary shaker. The serotype 2 classic Avery *Streptococcus pneumoniae* strain D39, *Streptococcus agalactiae* strain 2603V/R, and *E. coli* O157:H7 strain EDL933 were purchased from the ATCC. *S. pneumoniae* strain 0100993 is a type 3 human clinical isolate isolated by SmithKline Beecham. It was passaged through mice, recovered from the lung, and was the kind gift from Gee Lau in our department. There is no sequence for 0100993 available in the NCBI database. Its *nanS* homolog was cloned using primers designed for strain D39. An in-frame deletion of *E. coli* strain EDL933 *nanS* was constructed as previously described (11, 21), using the primer set shown in Table S1 in the supplemental material.

E. coli strains were grown in LB (Lennox formulation from Thermo Fisher) or M63 minimal medium (13), supplemented with 1 mg/ml diacetylated sialic acid (Neu5,9Ac_2) or 0.4% glycerol as a carbon source. *Streptococcus* spp. were grown on Columbia blood agar with 5% sheep blood (Remel) or in Todd-Hewitt broth (Remel) with 0.5% yeast extract. Cultures were incubated at 37°C in 5% CO_2.

Cloning *nanS* and its homologs or paralogs in pFC20K. Primers to clone each gene of interest into the HaloTag vector pFC20K were designed using the Flexi vector primer design tool on the Promega website (see Table S1 in the supplemental material). The genes to be cloned were amplified from boiled extracts or purified genomic DNA of the appropriate bacterial strain using PCR SuperMix high fidelity (Life Technologies), according to the manufacturer's protocol. The resulting PCR fragments were cloned into the HaloTag vector pFC20K (Promega), according to the manufacturer's instructions. Briefly, the PCR fragments were purified using the Wizard SV gel and PCR Clean-Up system (Promega) and treated with the appropriate restriction enzymes. The vectors were also digested with the restriction enzymes and then ligated to the digested PCR fragments. The ligation mixture was transformed into *E. coli* strain DH5α. Plasmid constructs were confirmed by restriction digestion of prepared minipreps (Wizard SV miniprep kit; Promega) with restriction enzyme sites in the pFC20K plasmid and the PCR product. The resulting plasmid in each case harbors the gene of interest with the coding sequence for the HaloTag fused to the gene's C-terminal end. Expression of the fusion is under the control of a T7 promoter.

PAGE analysis and purification of the HaloTag fusion proteins. Strain KRX harboring the pFC20K clones was grown overnight at 37°C in 5 ml of LB with 50 μg/ml kanamycin in a shaking water bath. The cultures were diluted 1:100 into fresh LB with 50 μg/ml kanamycin and grown to an A_{600} of approximately 0.4. At this point, 180 μl of the culture was removed and put into two small test tubes each, to which 20 μl of a 1:200 dilution of 6-carboxytetramethylrhodamine (TAMRA), a fluorescent dye which covalently binds to the HaloTag, was added in LB. One microliter of 20% rhamnose was added to one of the tubes to induce transcription of the T7 polymerase in KRX and thus transcription of the HaloTag fusion. The cultures were returned to the water bath for another 1.5 to 2 h. At the end of the incubation, the cells were pelleted and washed twice in 0.5 ml of phosphate-buffered saline (PBS). After the final spin, the cells were resuspended in 20 μl of PBS with 3.5 μl of protease inhibitor; 20 μl of PAGE sample buffer was added, and the sample was boiled for 5 min and centrifuged to remove debris. Samples of 20 μl were run on precast 4 to 20% Precise protein gels (Thermo). The gels were rinsed in distilled water after electrophoresis and the signals visualized on a Typhoon imager (GE Healthcare).

For purification of the HaloTag fusions, an overnight culture of KRX harboring the pFC20K clones was diluted 1:100 in 45 ml of LB with 50 μg/ml kanamycin. After 2 h of incubation, 20 ml of the culture was removed and put into a new flask, and 100 μl of 20% rhamnose was added. Growth of both cultures was continued for 1.5 h. Purification with the HaloLink resin was performed according to the manufacturer's directions, with minor modifications. Briefly, cells were pelleted and resuspended in 50 mM HEPES (pH 7.5) and 150 mM NaCl buffer with 500 μM EDTA, 1 mM dithiothreitol (DTT), and 0.005% IGEPAL detergent (Sigma). Cells were broken by sonication, after which debris was removed by centrifugation. The lysate was added to 2 ml of HaloLink and mixed for 1 h at room temperature on a tube rotator. The resin was pelleted, and the flowthrough was removed, saving a portion for PAGE. The resin was washed with the buffer described above, changing the NaCl concentration to 500 mM for the first wash, no NaCl for the second wash, and 150 mM for the third wash. Tobacco etch virus (TEV) protease (fused to a His tag) was added in the original buffer and incubated at room temperature on a rotator for 1 h, after which the resin was pelleted. The supernatant was removed and mixed with HisLink resin on the rotator for 20 min. The HisLink resin was pelleted, and the supernatant containing the purified protein was removed. After purification, the concentration of the purified esterase was determined using the Qubit 2.0 fluorometer (Invitrogen) and the Qubit protein assay kit, according to the manufacturer's directions.

Chromatographic analysis. Diacetylated sialic acid (Neu5,9Ac_2) and its enzymatically deacetylated product (Neu5Ac) were detected by high-performance liquid chromatography of samples fluorescently labeled with 1,2-diamino-4,5-methylenedioxybenzene (DMB). The analyses were carried out exactly as we previously described (11, 14, 15).

RNA isolation and quantitative PCR. Three individual cultures of *E. coli* O157:H7 strain EDL933 were grown overnight in M63 with glycerol as the carbon source. The cultures were diluted 1:100 in 2 fresh aliquots of M63 plus 0.4% glycerol each. After 4 h of growth at 37° with shaking, Neu5Ac (1 mg/ml final concentration) was added to one of each of the three cultures. When the cultures reached an A_{600} of 0.6, they were added to 2 volumes of RNAprotect bacteria (Qiagen), incubated for 10 min at room temperature, and then pelleted. RNA was purified from the pelleted cells using the Promega Wizard SV RNA isolation kit. When further DNase treatment was needed, it was done with Turbo DNase (Life Technologies). RNA from three of the *nanS* paralogs was quantified with the Qubit fluorometer and the Qubit RNA assay kit.

Quantitative PCR (qPCR) was performed using the Promega GoTaq 1-step reverse transcription-quantitative PCR (RT-qPCR) system, according to the manufacturer's protocol. Standard curves using serial dilutions of the HaloTag clones were done to ensure that the efficiency of each set of primers was close to 100%. The reaction parameters were 37°C for 15 min, 95°C for 10 min, and 40 cycles of 95°C for 10 s, 60°C for 30 s, and 72°C for 30 s, followed by a dissociation at 60°C to 95°C. The reactions were performed in an ABI7000 sequence detection system. The quantification cycle (C_q) values from the triplicate samples were averaged. Fold induction was calculated using the $R = 2^{-\Delta\Delta CT}$ method (16). The glucan biosynthesis protein gene *mdoG* was used as a reference gene for normalization (17). RNA of the Z5905 homolog of *E. coli* K-12 was analyzed nine times for the qPCR analysis.

Database search. The search for copies of *nanS* in all sequenced *Enterobacteriaceae* genomes to date (see Table S2 in the supplemental material) was conducted by searching the NCBI database prokaryotic genomes for any BioProjects with the relevant genera in the groups. The NanS primary structure from *E. coli* strain MG1655 was used in a BLAST search against all protein or translated nucleotide sequences for the species available in that genus. Some of these were complete genomes, and some were whole-genome shotgun projects. If no strains of a particular species contained an *nanS* homolog, only one accession number was provided, with a number indicating the number of other strains whose genomes are available, and these were also searched.

Esterase activity with a model substrate. The esterase activity of bacterial NanS was determined by monitoring the change in A_{450} per minute per milligram of protein, as previously described (14). *para*-nitrophenol acetate

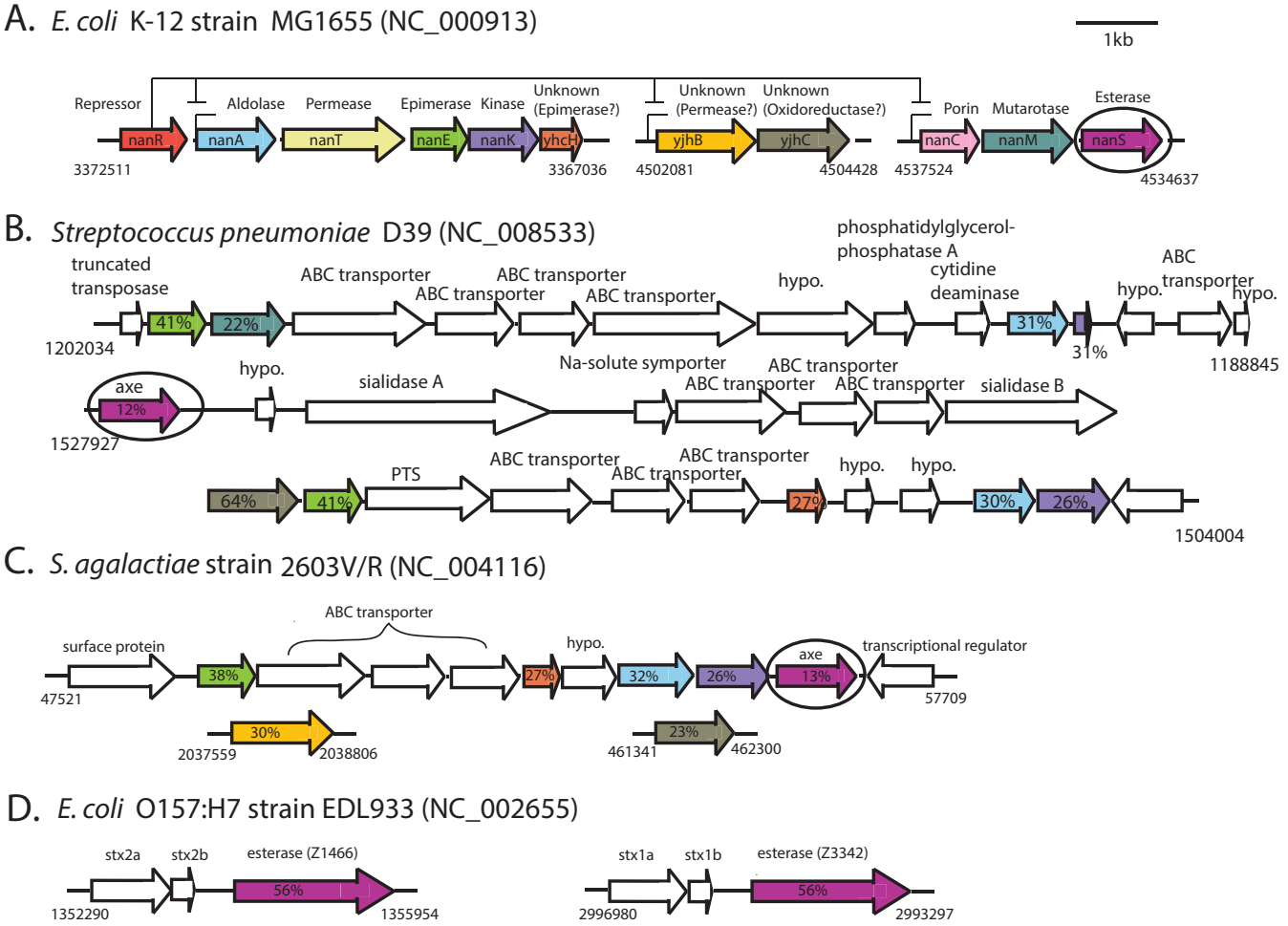

FIG 1 Genetic organization of *nan* genes in bacteria investigated in this study. Arrows indicate transcriptional direction and are drawn to scale. (A) Genes in the *E. coli* K-12 sialoregulon are color coded, with known or suspected functions listed above the genes, where the esterase gene *nanS* is circled. (B to D) Colored genes are homologs or paralogs of *E. coli* K-12 genes, with the percent identities of gene products indicated. Unshaded arrows indicate *nan* genes not present in *E. coli* K-12. Numbers underneath the lines indicate chromosomal distances in nucleotides from the origins. hypo., hypothetical protein; PTS, phosphotransferase system. Numbers in parentheses next to the strain names are accession numbers.

(pNP-Ac) was incubated at room temperature with extracts containing overproduced Halo-tagged polypeptides. The change in A_{450} was monitored using the kinetics program with a Beckman DU-640 spectrophotometer.

RESULTS AND DISCUSSION

Genetic organization of *nanS* in selected bacterial species or strains. We previously documented the wide variety of *nan* gene organizations in different bacterial commensals and pathogens (8). Most of these species, though not all, include one or more copies of *nanS*. However, *Salmonella enterica* serovar Typhimurium lacks *nanS* entirely and cannot use diacetylated Neu5Ac as a sole source of carbon due to a lack of the encoded *O*-acetyl esterase (11), documenting the essential function of NanS for the metabolism of *O*-acetylated sialic acids. We have commented at length on the evolutionary and functional reasons for the observed diversity of *nan* gene clusters (8, 18). Figure 1A shows the canonical *E. coli* K-12 *nanATEK-yhcH* operon required for catabolism of Neu5Ac, Neu5Gc, and probably all other sialic acid derivatives (1–3, 5, 12). The canonical operon is negatively regulated by NanR, as are the other two operons conferring coordinated control of the *E. coli* sialoregulon (12). Figure 1B and C shows that while *Streptococcus* species strains D39 and 2603V/R have a different *nan* organization from that of *E. coli* (Fig. 1A) and each other, they both include potential homologs of *E. coli nanS* with low primary structural identity to *E. coli* NanS. The low structural identity raises the question whether these so-called *axe* genes encode *O*-acetyl sialyl esterases or not.

As we have previously shown, enterohemorrhagic *E. coli* (EHEC) strains, such as O157:H7 strain EDL933 (Fig. 1D), include putative *nanS* paralogs located downstream of prophage encoding Shiga-like toxin 1 or 2 and 10 other paralogs in what appear to be cryptic lysogenic phage (8), as well as the canonical *nanS* located in a position identical to that in *E. coli* K-12 (Fig. 1A). No other known genes for the catabolism of sialic acid are located in the vicinity of the genes for the NanS paralogs in the prophage or prophage remnants. Since the canonical *nanS* (Z5905) of strain EDL933 has been noted to have some function in EHEC adherence (19), it was of interest to confirm the functions of the paralogs as encoding *O*-acetyl esterases. Four paralogs were chosen to

TABLE 1 Primary structural comparison of *E. coli* NanS to selected orthologs or paralogs in other species or strains

Organism (gene)	Accession no.	Mol wt	% identity to:	
			NanS[a]	Axe[b]
E. coli K-12 (*nanS*)	NP_418729	36,878	100	18
S. pneumoniae (SPD_1506)	ABJ55220	37,439	12	12
S. agalactiae (gbs0040)	NP_734510	37,225	14	12
E. coli O157 (Z5905)[c]	NP_290925	36,886	97	17
E. coli O157 (Z1466)	NP_286978	68,619	55	22
E. coli O157 (Z2108)	NP_287566	34,432	46	15
E. coli O157 (Z3927)	NP_289184	42,225	56	22
E. coli O157 (Z3342)	NP_288671	68,776	55	23

[a] Percent identity of the DUF303 domains from each protein (represented by white rectangles in Fig. 2) to K-12 NanS.
[b] Percent identity to *Rhodopirellula baltica* SH1 AxeA (GenBank accession no. NP_866453).
[c] Z5905 encodes the predicted NanS that is present in *E. coli* K-12.

be representative of the different sizes of the polypeptides due to additional N-or C-terminal domains. Table 1 summarizes the species and strains investigated in this study and the relative sizes and percent identities of their *nanS* homologs or paralogs to canonical *E. coli* NanS.

The reason for the size variation of paralogous *nanS* gene products shown in Table 1 was identified by aligning their predicted primary structures to canonical NanS (Fig. 2). While most polypeptides included the four blocks typical of esterases, EHEC copies had N- or C-terminal domains of unknown function fused to NanS. The N-terminal domains (most often DUF1737) have no known function and are found exclusively in prophage, cryptic prophage, or prophage remnants associated with *nanS* (DUF303). The C-terminal domains have no assigned DUF number but are also associated with prophage and cryptic prophage. One (Z2108) had a deletion of block IV and another (Z3927) a 6-amino-acid addition at the NanS C terminus. Neither of these two included long C-terminal domains, thus accounting for their lower predicted molecular weights (Table 1).

EHEC *nanS* is not regulated by NanR. As shown in Fig. 1A, the *E. coli* sialoregulon that includes *nanS* is coordinately controlled by the NanR repressor. The inducer of this regulon, Neu5Ac, upon binding shifts the quaternary structure of NanR from functional homodimer to monomers (12). The NanR binding site is not present for EHEC *nanS* paralogs, suggesting that these paralogs are not affected by the inducer of the sialoregulon. This expectation was demonstrated by qPCR (Fig. 3), which shows the expected induction of the canonical *nanS* gene (Z5905), while the three different EHEC paralogs indicated were not induced. These results show that EHEC *nanS* paralogs are not part of the sialoregulon, strongly suggesting that the putative esterases encoded by these genes do not function in any significant manner to augment growth on diacetylated sialic acids. The implication is that the primary function of the multiple EHEC NanS paralogs is something other than catabolism, although as shown below, they can compensate for a loss of canonical *nanS* for growth on Neu5,9Ac$_2$ when Z5905 is deleted from EHEC strains.

Purification and biochemical characterization of *E. coli* K-12 NanS paralogs. To facilitate biochemical analysis of NanS, genes were cloned as C-terminal fusions to HaloTag, a 33-kDa polypeptide that binds TAMRA, allowing sensitive fluorometric detection of tagged polypeptides. Figure 4 shows the results of various gene products overproduced in strain KRX (see Materials and Methods). This method worked well for most tagged polypeptides, except Z3342, which for unknown reasons was not detectable. Note that extraneous signals result from the high sensitivity of TAMRA labeling, indicating the detection of probably proteolytic by-products that, as shown below, are not detectable in purified samples.

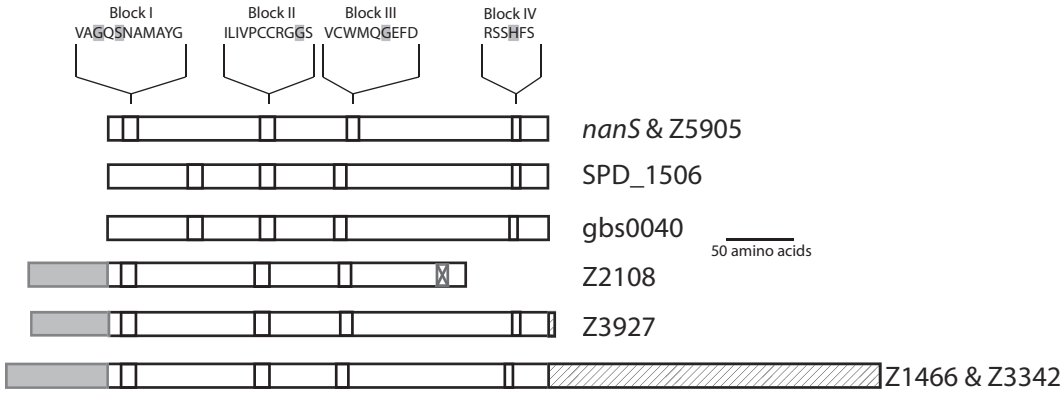

FIG 2 Schematic alignment of esterase primary structures of *nanS* homologs and paralogs investigated in this study. The polypeptides encoded by the genes indicated were aligned using the ClustalW multiple-sequence alignment algorithm in the MacVector software. The location in each polypeptide of the esterase-conserved blocks I to IV (14) is indicated by boxes within the bars. The gray-shaded regions represent similar amino-terminal domains, and the hatched areas represent differing carboxy-terminal domains not present in NanS or Z5905. The box with the X in Z2108 indicates an 8-amino-acid deletion in this protein compared to NanS when the proteins sequences are aligned. This polypeptide was not further investigated.

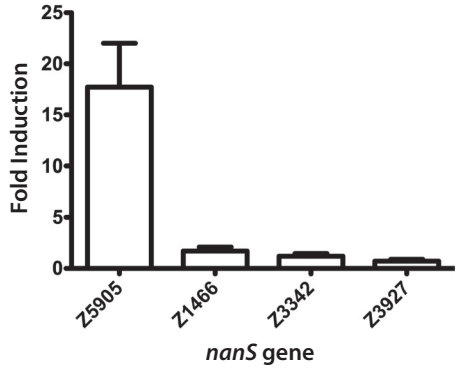

FIG 3 EHEC *nanS* paralogs are not induced by Neu5Ac. Induction of EHEC strain EDL933 gene Z5905, homologous to the canonical *E. coli* K-12 *nanS*, and prophage paralogs indicated were analyzed by qPCR (see Materials and Methods). Nine independent samples of Z5905 were analyzed, and three samples each were analyzed for paralogs. Note that a fold induction of 1.0 indicates no induction. Data were statistically analyzed by GraphPad Prism software to indicate standard deviations.

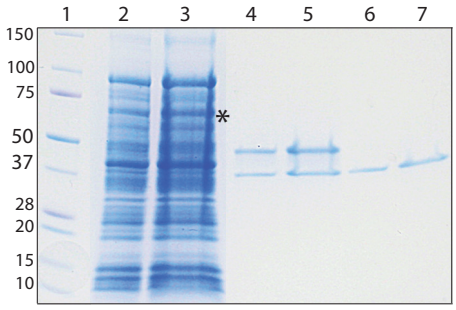

FIG 5 Purification of HaloTag *E. coli* K-12 NanS. *E. coli* K-12 NanS (37 kDa) was overproduced as a C-terminal HaloTag fusion (33 kDa) by induction in strain KRX with rhamnose (see Materials and Methods). Lanes 2 and 3 show successive 10- and 20-μl loadings of soluble proteins from induced lysate. The asterisk indicates the position of the 70-kDa fusion polypeptide. Extract was added to affinity-tag resin, washed extensively, and the product released by hydrolysis with TEV protease (lanes 4 and 5), showing the 37-kDa esterase and 50-kDa protease. Lanes 6 and 7 show purification of the esterase after the removal of TEV by nickel-affinity chromatography. The sample in lane 6 represents 1/70 of the total purified protein from a 25-ml induced culture. Lane 1 shows markers with sizes in kilodaltons at the left. The gel was stained with Coomassie blue.

The sizes of overproduced polypeptides are consistent with those shown in Table 1 plus the HaloTag, except for Z2108, which as noted, was not investigated further. The band in Fig. 4 representing the overproduced Z2108 plus HaloTag is likely a proteolytic product of the fusion polypeptide. Z3927 with the HaloTag runs as an approximately 75-kDa polypeptide, Z1466 with the HaloTag as a 102-kDa polypeptide, and SPD_1506 as a 71-kDa polypeptide.

Figure 5 shows an example of the purification scheme with amplified NanS from *E. coli* K-12. Lanes 2 and 3 show different loadings of overproduced NanS (37 kDa) fused to the 33-kDa HaloTag. The asterisk in Fig. 5 identifies the 70-kDa polypeptide in the crude extract of induced KRX, which is not detectable in the uninduced extract (not shown). The induced extract was loaded onto HaloTag resin, washed, and eluted after cleavage with TEV protease, resulting in 37-kDa NanS and 50-kDa TEV (Fig. 5, lanes 4 and 5). The protease is removed by nickel-affinity chromatography, resulting in purified NanS (Fig. 5, lanes 6 and 7).

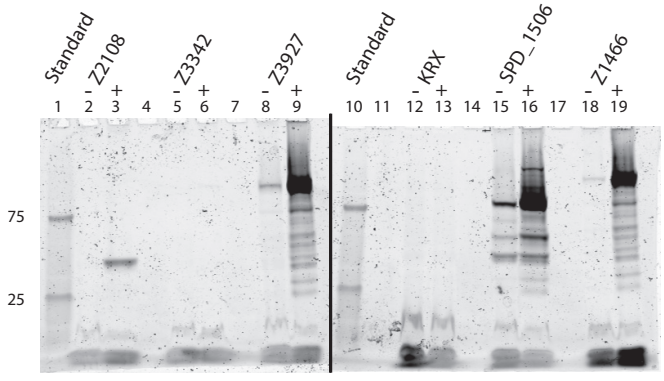

FIG 4 Overproduction of HaloTag *nanS* fusions. HaloTag plasmids harboring various *nanS* homologs or paralogs were expressed in strain KRX and induced or not with rhamnose, as described in Materials and Methods. Cells were labeled with TAMRA and extracts analyzed after PAGE. Molecular size markers (lanes 1 and 10) were included, with the molecular mass in kilodaltons shown at left. The vertical line indicates the alignment of two separate gels. Lanes 4, 7, 11, 14, and 17 are blank. Lanes 2, 5, 8, 12, 15, and 18 show uninduced extracts (−) from plasmids expressing tagged Z2108, Z3342, Z3927, KRX with no plasmid, D39 SPD_1506, and Z1466, respectively. The respective induced extracts (+) are shown in lanes 3, 6, 9, 13, 16, and 18.

To demonstrate the esterase activity of the purified *E. coli* K-12 NanS, either purified enzyme (1 ng) or 5 μl of induced extract was incubated with Neu5,9Ac_2 and the activity assayed by monitoring its conversion to Neu5Ac. As demonstrated previously (11), the substrate is contaminated by small amounts of free Neu5Ac (Sia) and the Neu5,8Ac_2 (5, 8) and Neu5,7Ac_2 (5, 7) diacetylated forms (Fig. 6). The presence of 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo) and pyruvate (Pyr) contaminants in crude extracts has been previously described (12). When either the purified or crude extract from HaloTag *nanS* plasmid transformed into the *E. coli* Δ*nanS* mutant was inactivated by preheating samples at 95°C for 5 min, the chromatographic profiles were identical to the substrate profile (Fig. 6), indicating the inactivation of NanS and thus supporting the expected biochemical activity of the esterase, demonstrating its relative activity against the various diacetylated products and the subsequent increase in the Sia peak. The activity of the induced extract was sufficiently great to collapse nearly all diacetylated substrates to Sia. The combined results indicate that HaloTag fusions can be used for analysis of esterase activity without the need for time-consuming affinity chromatography steps, and that crude extracts can be used to detect esterase activity.

Sialate esterase activity of *S. pneumoniae* Axe (NanS). There are reports that the catabolism of sialic acid is important for *S. pneumoniae* colonization (8). Despite the low primary structural identity of *S. pneumoniae* Axe to *E. coli* NanS, the location of *axe* near *S. pneumoniae nan* genes suggested that *axe* encodes a sialate *O*-acetyl esterase (Fig. 1). To confirm this suspected function, we transformed an *E. coli* K-12 *nanS* deletion strain (11) with HaloTag fusions of *axe* from *S. pneumoniae* strains D39 and 0100933 and *E. coli* K-12 *nanS*. Extracts were incubated with Neu5,9Ac_2 and subjected to chromatographic analysis. Figure 7A shows the expected profile of substrate with no extract. Figure 7A to C shows a nearly complete collapse of the *O*-acetylated substrate to Neu5Ac (Sia), demonstrating that *S. pneumoniae axe* does in fact code for NanS. This result was further confirmed by complementation for growth on Neu5,9Ac_2 as a sole source of

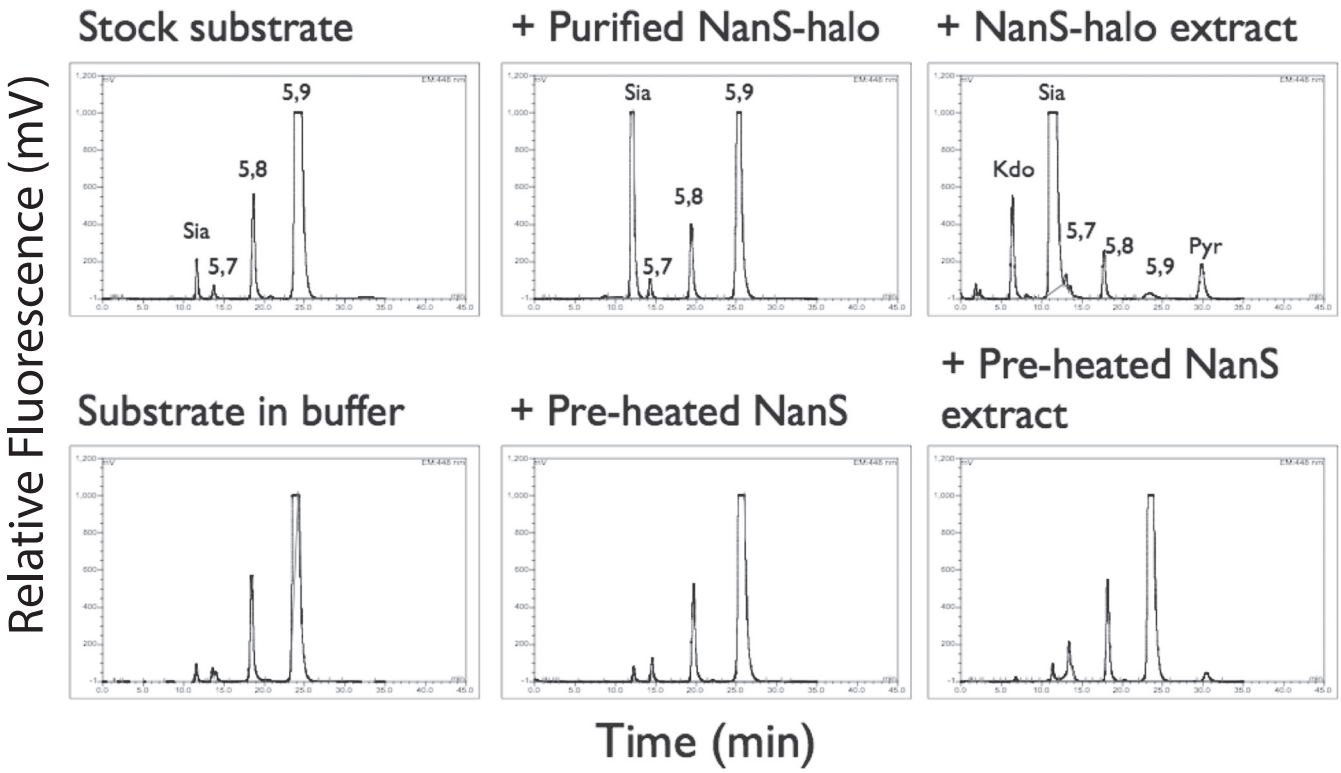

FIG 6 DMB analysis of purified *E. coli* K-12 NanS. Samples containing 1 ng of purified NanS (Fig. 5) or 5 μl of crude extract were incubated with Neu5,9Ac$_2$ for 1 h and then analyzed by fluorescent chromatography of DMB-derivatized samples. See the text for details.

carbon, where the *E. coli* K-12 *nanS* deletion strain grew to an A_{600} of 0.06 after overnight incubation. In contrast, the strain harboring the canonical *E. coli nanS* plasmid grew to an A_{600} of 0.40, while the strain harboring the D39 *axe* grew to an A_{600} of 0.87, and two independent clones of strain 0100933 *axe* each grew to an A_{600} of 0.91, confirming that *axe* is in fact a functional homolog of *nanS*. Indeed, the more robust complementation by the *S. pneumoniae* clones implies that this Axe may be a more active esterase than NanS. This conclusion was suggested by an assay of induced extracts of strain D39 *axe* and paralog Z1466 (see Table S3 in the supplemental material). However, we cannot exclude that the complementation and assay results are not a result of differences in gene expression. The results strongly suggest that all *nanS* or *axe* homologs located near canonical *nan* genes in bacterial commensals and pathogens are sialate *O*-acetyl esterases functioning in the metabolism of *O*-acetylated sialic acids. For reasons unknown, expression of the cloned *S. agalactiae axe* homolog was undetectable, although its similarity to *S. pneumoniae axe* suggests that it too encodes a functional NanS.

Sialate esterase activity of EHEC prophage-encoded NanS. The results of qPCR analysis (Fig. 3) indicated that EHEC *nanS* paralogs were constitutively expressed. To demonstrate this experimentally, we constructed a deletion of the canonical *nanS* in strain EDL933, which is identical to that of *E. coli* K-12 (Fig. 1A). In the absence of a carbon source, the A_{600} of the deletion strain after overnight incubation in minimal medium was 0.07. In con-

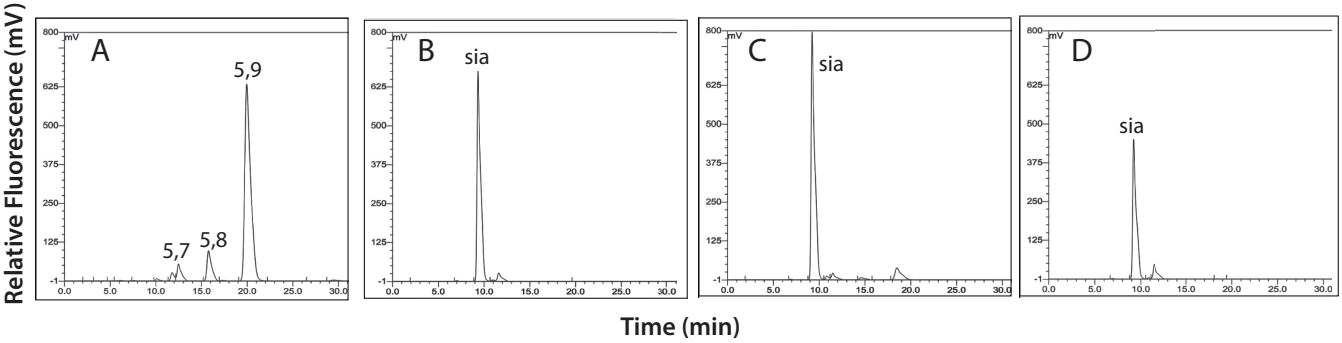

FIG 7 DMB analysis of *S. pneumoniae* Axe. Tagged *S. pneumoniae* strain D39 or 0100993 *axe* or *E. coli* K-12 *nanS* plasmids were transformed into the *E. coli* K-12 *nanS* deletion strain. Extracts were incubated for 4 h with Neu5,9Ac$_2$ and subjected to chromatographic analysis of DMB-labeled samples. (A to D) Substrate with no extract (A), incubated with NanS (B), incubated with D39 (C), and incubated with 0100993 (D).

trast, wild-type EDL933 grew to an A_{600} of 0.50 on Neu5,9Ac$_2$, while the deletion strain grew to an A_{600} of 0.44, confirming that the phage-encoded paralog(s) can compensate for a loss of canonical NanS. In a related experiment, HaloTag plasmids with fusions to Z1466, Z3342, and Z3927 (Fig. 1D and 2) transformed into the *E. coli* K-12 *nanS* deletion strain grew to an A_{600} of 0.70, 0.50, and 0.44, respectively, demonstrating complementation of the canonical *nanS* deletion. It is possible that while the cloned genes were induced in a similar manner, differences in relative degradation and folding account for the differences in growth.

Conclusions. Among dozens of sequenced genomes from *Enterobacteriaceae* genera, *E. coli* is one of the few species to include *nanS* (see Table S2 in the supplemental material), supporting our previous suggestion that the expression of sialate esterase by *E. coli* in part accounts for its success as a mucosal commensal (8). In contrast, most other mucosal commensals and pathogens include homologs of *nanS*, supporting an important function of esterase in the host-bacterium interaction (8). The results presented here showing that *S. pneumoniae axe* is in fact a functional *nanS* homolog support these suppositions, as does the demonstration of a functional NanS from the oral pathogen *Tannerella forsythia*, a member of the *Bacteroidetes* group (20). We also previously documented that EHEC strains contain up to 12 *nanS* paralogs and the one canonical *nanS* copy (8). Some, as shown in Fig. 1D, are part of intact prophage, whereas most are part of phage chromosomal remnants that probably resulted from past recombinational events. Our current results show that at least some of these paralogs encode active esterases. However, the highly varied occurrence of N- or C-terminal fusions of *nanS* to domains of unknown function in genes located immediately downstream of *stx* genes is curious (Fig. 1D). Our results indicate that the function of these *nanS* paralogs is not to augment sialometabolism, although they can compensate for a loss of canonical *nanS*. A signature tag mutagenesis study of strain EDL933 identified a function of Z5905 in the colonization of the bovine intestinal tract (19). This would seem to exclude any function of the prophage *nanS* paralogs. However, if prophage-encoded NanS is excreted, as our preliminary results suggest, despite the lack of an obvious signal sequence, in combination with the domains of unknown function, the esterase could function in remodeling the host mucosal surface to enhance the bacterium-host association. However, the location of these prophage paralogs just downstream of the *stx* genes also suggests the function of prophage-encoded NanS could be to augment toxin access to the host cell surface. Our combined results present numerous experimental avenues to better understand the functions of sialobiology in host-microbe interactions.

ACKNOWLEDGMENTS

We thank Jamie L. Jirik and Bahaa Fadl-Alla for technical assistance and Eric Sixmister and Gee Lau for helpful discussions and sharing of strains. We are grateful to Kerry Helms for his expert assistance with graphics and preparation of illustrations.

FUNDING INFORMATION

This work, including the efforts of Susan M. Steenbergen and Eric R. Vimr, was funded by HHS | NIH | National Institute of Allergy and Infectious Diseases (NIAID) (AI042015).

REFERENCES

1. **Vimr ER, Troy FA.** 1985. Identification of an inducible catabolic system for sialic acids (*nan*) in *Escherichia coli*. J Bacteriol **164:**845–853.
2. **Vimr ER, Troy FA.** 1985. Regulation of sialic acid metabolism in *Escherichia coli*: role of *N*-acylneuraminate pyruvate-lyase. J Bacteriol **164:**854–860.
3. **Plumbridge J, Vimr E.** 1999. Convergent pathways for the utilization of the amino sugars *N*-acetylglucosamine, *N*-acetylmannosamine and *N*-acetylneuraminic acid by *Escherichia coli*. J Bacteriol **181:**47–54.
4. **Vimr E, Lichtensteiger C, Steenbergen S.** 2000. Sialic acid metabolism's dual function in *Haemophilus influenzae*. Mol Microbiol **36:**1113–1123. http://dx.doi.org/10.1046/j.1365-2958.2000.01925.x.
5. **Kalivoda KA, Steenbergen SM, Vimr ER, Plumbridge J.** 2003. Regulation of sialic acid catabolism by the DNA binding protein NanR in *Escherichia coli*. J Bacteriol **185:**4806–4815. http://dx.doi.org/10.1128/JB.185.16.4806-4815.2003.
6. **Steenbergen SM, Lichtensteiger CA, Caughlan R, Garfinkle J, Fuller TE, Vimr ER.** 2005. Sialic acid metabolism and systemic pasteurellosis. Infect Immun **73:**1284–1294. http://dx.doi.org/10.1128/IAI.73.3.1284-1294.2005.
7. **Tatum FM, Tabatabai LB, Briggs RE.** 2009. Sialic acid uptake is necessary for virulence of *Pasteurella multocida* in turkeys. Microb Pathog **46:**337–344. http://dx.doi.org/10.1016/j.micpath.2009.04.003.
8. **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.
9. **Vimr ER, Steenbergen SM.** 2006. Targeting microbial sialic acid metabolism for new drug development, p 125–150. *In* Bewley CA (ed), Protein-carbohydrate interactions in infectious disease. Royal Society of Chemistry, London, United Kingdom.
10. **Johnston JW, Apicella MA.** 2008. Sialic acid metabolism and regulation by *Haemophilus influenzae*: potential novel antimicrobial therapies. Curr Infect Dis Rep **10:**83–84. http://dx.doi.org/10.1007/s11908-008-0014-y.
11. **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.
12. **Kalivoda KA, Steenbergen SM, Vimr ER.** 2013. Control of the *Escherichia coli* sialoregulon by transcriptional repressor NanR. J Bacteriol **195:**4689–4701. http://dx.doi.org/10.1128/JB.00692-13.
13. **Miller JH.** 1972. Experiments in molecular genetics. Cold Spring Harbor Press, Cold Spring Harbor, NY.
14. **Steenbergen SM, Lee Y-C, Vann WF, Vionnet J, Wright LF, Vimr ER.** 2006. Separate pathways for *O*-acetylation of polymeric and monomeric sialic acids and identification of sialyl *O*-acetyl esterase in *Escherichia coli* K1. J Bacteriol **188:**6195–6206. http://dx.doi.org/10.1128/JB.00466-06.
15. **Steenbergen SM, Vimr ER.** 2013. Chromatographic analysis of the *Escherichia coli* polysialic acid capsule. Methods Mol Biol **966:**109–120. http://dx.doi.org/10.1007/978-1-62703-245-2_7.
16. **Livak KJ, Schmittgen TD.** 2001. Analysis of relative gene expression data using real-time quantitative PCR and $2^{-\Delta\Delta CT}$ method. Methods **25:**402–408. http://dx.doi.org/10.1006/meth.2001.1262.
17. **Heng SS, Chan OY, Keng BM, Ling MH.** 2011. Glucan biosynthesis protein G is a suitable reference gene in *Escherichia coli* K-12. ISRN Microbiol **2011:**469053. http://dx.doi.org/10.5402/2011/469053.
18. **Vimr ER, Kalivoda KA, Deszo EL, Steenbergen SM.** 2004. Diversity of microbial sialic acid metabolism. Microbiol Mol Biol Rev **68:**132–153. http://dx.doi.org/10.1128/MMBR.68.1.132-153.2004.
19. **Dziva F, van Diemen PM, Stevens MP, Smith AJ, Wallis TS.** 2004. Identification of *Escherichia coli* O157:H7 genes influencing colonization of the bovine gastrointestinal tract using signature-tagged mutagenesis. Microbiology **150:**3631–3645. http://dx.doi.org/10.1099/mic.0.27448-0.
20. **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.
21. **Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H.** 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol **2:**2006.0008.