ISSN 0006-2979, Biochemistry (Moscow), 2015, Vol. 80, No. 7, pp. 901-914. © Pleiades Publishing, Ltd., 2015.
Published in Russian in Biokhimiya, 2015, Vol. 80, No. 7, pp. 1072-1087.

REVIEW

O-Antigen Modifications Providing Antigenic Diversity of *Shigella flexneri* and Underlying Genetic Mechanisms

Y. A. Knirel[1]*, Qiangzheng Sun[2], S. N. Senchenkova[1], A. V. Perepelov[1], A. S. Shashkov[1], and Jianguo Xu[2]

[1]*Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, 119991 Moscow, Russia; fax: (499) 137-6148; E-mail: yknirel@gmail.com*
[2]*State Key Laboratory for Infectious Disease Prevention and Control, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, P.O. Box 5, Changping, Beijing 102206, China; E-mail: sunqiangzheng@icdc.cn; xujianguo@icdc.cn*

Received December 11, 2014
Revision received February 24, 2015

Abstract—O-Antigens (O-specific polysaccharides) of *Shigella flexneri*, a primary cause of shigellosis, are distinguished by a wide diversity of chemical modifications following the oligosaccharide O-unit assembly. The present review is devoted to structural, serological, and genetic aspects of these modifications, including O-acetylation and phosphorylation with phosphoethanolamine that have been identified recently. The modifications confer the host with specific immunodeterminants (O-factors or O-antigen epitopes), which accounts for the antigenic diversity of *S. flexneri* considered as a virulence factor of the pathogen. Totally, 30 O-antigen variants have been recognized in these bacteria, the corresponding O-factors characterized using specific antibodies, and a significant extension of the serotyping scheme of *S. flexneri* on this basis is suggested. Multiple genes responsible for the O-antigen modifications and the resultant serotype conversions of *S. flexneri* have been identified. The genetic mechanisms of the O-antigen diversification by acquisition of mobile genetic elements, including prophages and plasmids, followed occasionally by gene mobilization and inactivation have been revealed. These findings further our understanding of the genetics and antigenicity of *S. flexneri* and assist control of shigellosis.

DOI: 10.1134/S0006297915070093

Key words: *Shigella flexneri*, O-antigen, O-polysaccharide structure, serotype-converting bacteriophage, transposon, plasmid, serotyping, immunodeterminant

Shigellosis, or bacillary dysentery, is an acute diarrheal disease that remains an important public health challenge, especially in developing countries. There are an estimated about 164.7 million shigellosis cases annually worldwide causing 1.1 million deaths, with the majority involving children under five years old [1]. The causative agent of shigellosis is *Shigella* spp., nonmotile, nonspore-forming facultative anaerobic Gram-negative bacteria, which are among the bacterial pathogens most frequently isolated from patients with diarrhea. Invasion by these bacteria of the colonic and rectal mucosa and the following inflammatory response provoke massive mucosal destruction reflected in strong abdominal cramps and stools containing blood and mucus [2]. Based on biochemical properties and O-antigen specificity, the genus *Shigella* is divided into four species or subgroups, including *S. dysenteriae*, *S. flexneri*, *S. boydii*, and *S. sonnei* [3], although genetically they all but *S. boydii* type 13 are clones of *Escherichia coli* [4]. *Shigella flexneri* is the predominant species causing shigellosis in developing countries and the second, after *S. sonnei*, most common in industrialized countries [5-7].

Being a serologically heterogeneous species, *S. flexneri* is further divided into various serotypes and subtypes. A commercially available monovalent antisera kit (Denka Seiken, Japan) and monoclonal antibody reagents (MASFs) (Reagensia AB, Sweden) are widely used for serotyping *S. flexneri* isolates. The serospecificity of *S. flexneri* is defined by a combination of immunodeterminants (O-factors), which reside on the O-antigen located on the bacterial cell surface [3]. The O-antigen, also called O-specific polysaccharide or O-polysaccha-

Abbreviations: GalA, galacturonic acid; GalNAc, 2-acetamido-2-deoxygalactose; GlcNAc, 2-acetamido-2-deoxyglucose; IS, insertion sequence; LPS, lipopolysaccharide; PEtN, phosphoethanolamine; Rha, rhamnose; Sf, *Shigella flexneri*.

* To whom correspondence should be addressed.

ride, is a part of the outer-membrane lipopolysaccharide (LPS) and is linked to the lipid moiety (lipid A) via a core oligosaccharide. The O-polysaccharide consists of many oligosaccharide repeats (O-units). Structure and serology of the O-antigens of *S. flexneri* have been intensively studied for the last 50 years ([8-10] and references therein).

The O-antigens of *S. flexneri* are synthesized by the O-antigen polymerase (Wzy)/flippase (Wzx)-dependent pathway, whereby the O-unit is preassembled on a lipid carrier on the cytoplasmic side of the inner membrane, and after translocation (flipping) to the periplasmic side mediated by Wzx, is polymerized by Wzy with participation of a chain length regulator Wzz. There are two basal *S. flexneri* O-polysaccharide structures and, correspondingly, two non-variable gene clusters for O-antigen biosynthesis: one for serotype 6 and the other for the remaining serotypes. The bacteria of the two groups have different evolutionary origins and belong to different lineages of *Shigella* clones of *E. coli* [4]. As in most *E. coli* clones including the other *Shigella* spp., the O-antigen gene cluster maps between the housekeeping genes *galF* and *gnd* on the chromosome [11]. It contains genes for synthesis of a nucleotide (deoxythymidine diphosphate) precursor of L-rhamnose, a specific monosaccharide component of both basal O-polysaccharides, for glycosyltransferases necessary for the assembly of the O-unit, and O-antigen processing genes: *wzx* for flippase and *wzy* for O-antigen polymerase.

Molecular typing targeting specific genes in the O-antigen cluster, including *wzx*, *wzy*, and glycosyltransferases genes, detects all *S. flexneri* non-6 serotypes as a single group [12]. *Escherichia coli* O13, O129, and O135 having the same basal O-antigen structure [10, 13] and essentially identical O-antigen gene cluster [11] also fall in this group, and, similarly, *E. coli* O147 forms one molecular group and shares the O-antigen structure with *S. flexneri* serotype 6 [11, 12]. The closely related *E. coli* and *S. flexneri* clones can be differentiated using PCR assays based on other genes not related to the O-antigen synthesis [12].

The O-antigens of *S. flexneri* non-6 serotypes are highly diverse due to various chemical modifications to the basal structure giving rise to the observed serological heterogeneity [3]. A number of genes outside the O-antigen cluster are involved in the modifications [14], which occur after the O-unit assembly and before the transfer of the mature O-polysaccharide to the lipid A-core region of the LPS. A molecular approach targeting specific O-antigen modification genes identified by that time has been developed for serotyping *S. flexneri* within the group of non-6 serotypes [15].

The O-antigen plays an important role in the pathogenesis of *S. flexneri*; particularly, it protects the bacteria from the lytic action of serum complement and promotes adherence and internalization of bacteria to intestinal epithelial cells [16-18]. Creating antigenic diversity by O-antigen modifications is considered as an important virulence factor of *S. flexneri* that enhances survival of the pathogens because the host has to mount a specific immune response to each serotype [19]. Moreover, such modification as glucosylation at certain sites promotes invasion of *S. flexneri* into host cells mediated by the type III secretion system [16].

In previous reviews devoted to *S. flexneri* O-antigens [9, 14, 19], modifications known by that time, including glucosylation at various sites and O-acetylation at one site (on Rha^{I}), have been considered in detail. Recently, more sites of O-acetylation [10, 13, 20-25] and a novel modification type, phosphorylation with phosphoethanolamine (PEtN) [26-29], have been identified and genetic bases of the new modifications have been elucidated. The present review summarizes structural, serological, and genetic aspects of the O-antigen modifications in *S. flexneri* with emphasis on the new findings.

STRUCTURES AND IMMUNOSPECIFICITY OF THE O-ANTIGENS

Except for serotype 6, all known *S. flexneri* serotypes (1-5, 7, X, Y) share the O-polysaccharide backbone (1) composed of tetrasaccharide O-units containing three L-rhamnose residues (Rha^{I}-Rha^{III}) and one residue of 2-acetamido-2-deoxy-D-glucose (GlcNAc) [30].

$$\rightarrow 2)\text{-}\alpha\text{-L-Rha}p^{III}\text{-}(1\rightarrow 2)\text{-}\alpha\text{-L-Rha}p^{II}\text{-}(1\rightarrow 3)\text{-}$$

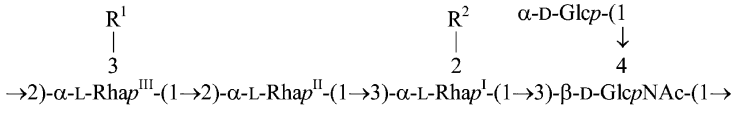
$$\text{-}\alpha\text{-L-Rha}p^{I}\text{-}(1\rightarrow 3)\text{-}\beta\text{-D-Glc}p\text{NAc-}(1\rightarrow \qquad (1)$$

The polysaccharide (1) present in serotype Y is characterized by two antigenic specificities labeled dual group O-factor 3,4. A structural domain that defines this O-factor has not been completely identified yet [31, 32]. In some cases, its manifestation is ambiguous as strains otherwise identical in the O-antigen structure and the presence of other immunodeterminants may express or may not express O-factor 3,4 (e.g. former serotypes 3b and 3c, which have been proposed to be combined into one serotype 3b [10]). The polysaccharide (1) can be modified by adding various chemical groups (α-D-glucopyranosyl, O-acetyl, phosphoethanolamine) to different sugars giving rise to enormously diverse O-antigen structures and, correspondingly, to serological heterogeneity, which is the basis for serotyping of *S. flexneri* strains (table).

Glucosylation may occur on any monosaccharide in the polysaccharide (1) giving rise to type O-factors I, II, IV, and V when present at various positions on Rha^{I}, Rha^{II}, or GlcNAc, and to dual group O-factor 7,8 when present on Rha^{III} (Fig. 1). The type O-factors define serotypes 1, 2, 4, and 5, respectively, whereas the group O-factor 7,8 may be expressed in different serotypes and occur in combination with various type O-factors [8, 9, 34, 35] (table). As a result, in the O-units of some subtypes

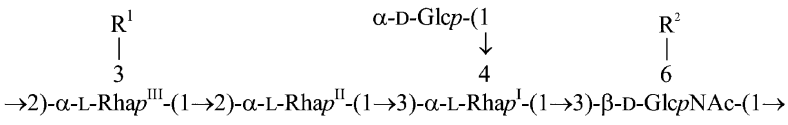
Structures of the O-polysaccharides of *S. flexneri*. Included are both serotypes already approved internationally and provisional serotypes that express epitopes associated with newly identified O-acetyl and PEtN groups

Serotype 1



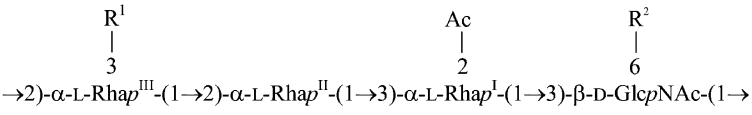
subtype	antigenic formula	R^1	R^2	reference
1a	I: -	H	H	[8, 33]
$1a_1$	I: 9	Ac	H	[21]
1b	I: 6; 9	Ac	Ac	[21]
1d	I: 7,8	α-D-Glc*p*	H	[34]

Serotype 2



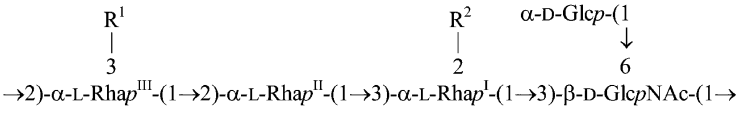
subtype	antigenic formula	R^1	R^2	reference
$2a_1$	II: 10	H	Ac	[33]
$2a_2$	II: 9; 10	Ac	Ac	[20, 21]
2b	II: 7,8	α-D-Glc*p*	H	[8]
$2b_1$	II: 7,8; 10	α-D-Glc*p*	Ac	[a]

Serotype 3



subtype	antigenic formula	R^1	R^2	reference
3a	III: 6; 7,8	α-D-Glc*p*	H	[8] [b]
$3a_1$	III: 6; 7,8; 10	α-D-Glc*p*	Ac	[10]
3b	III: 6	H	H	[8]

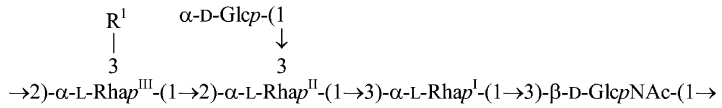
Serotype 4



subtype	antigenic formula	R^1	R^2	reference
4a	IV: –	H	H	[8]
4av	IV: IV-1	PEtN	H	[26, 27]
4b	IV: 6	H	Ac	[8]

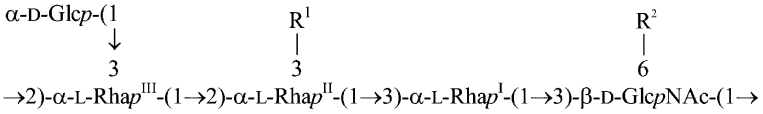
Table (Contd.)

5



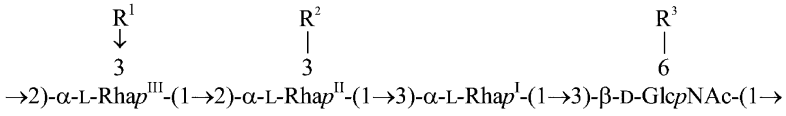
subtype	antigenic formula	R^1	reference
5a	V: –	H	[33, 35]
$5a_1$	V: 9	Ac	[13]
5b	V: 7,8	α-D-Glc*p*	[35]

X



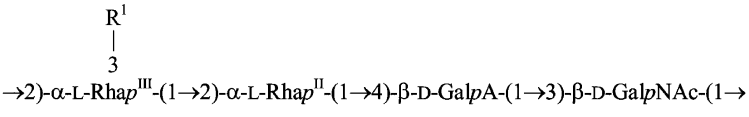
subtype	antigenic formula	R^1	R^2	reference
X	–: 7,8	H	H	[35]
X_1	–: 7,8; 10	H	Ac	[c]
Xv	–: 7,8; IV-1	PEtN	H	[27]

Y



subtype	antigenic formula	R^1	R^2	R^3	reference
Y	–: 3,4	H	H	H	[28, 30]
Y_1	–: 9	Ac	H	H	[d]
Y_2	–: 9; 10	Ac	H	Ac	[10]
Yv	–: IV-1	PEtN	PEtN	H	[e]
Yv_1	–: IV-1; 10	PEtN	PEtN	Ac	[28, 29]

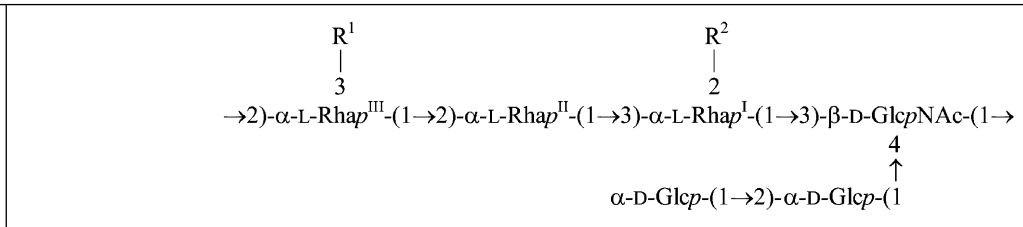
6



subtype	antigenic formula	R^1	reference
6	VI: 9	Ac	[10, 36]
6v[f]			

Table (Contd.)

7



subtype	antigenic formula	R^1	R^2	reference
7a (1c)	VII (IC): –	H	H	[37]
$7a_1$	VII: 9	Ac	H	[25]
7b	VII: 6	H	Ac	[38]

Note: Rha^{III} and GlcNAc are O-acetylated non-stoichiometrically. A minor 4-O-acetylation on Rha^{III} that occurs alternatively to the major 3-O-acetylation on Rha^{III} is not shown. In the antigenic formulae, type and group O-factors are indicated before and after colon, respectively. O-factor 3,4 associated with the O-polysaccharide backbone is variably expressed and, except for serotype Y, is omitted from the antigenic formulae in the table.

[a-e] Authors' unpublished data.

[a] Strain 2005122; the degree of 6-O-acetylation on GlcNAc ~75%.

[b] Strain 2001019.

[c] Strain 2005128; the degree of 6-O-acetylation on GlcNAc ~75%.

[d] Strain 06AH74; the degree of 3/4-O-acetylation on Rha^{III} ~40/25%.

[e] Strains 06HN054 and 06HN303.

[f] Proposed based on the reactivity with MASF IV-1 [45, 46] with no data on the O-antigen structure and genetics available.

there are two side-chain glucosyl groups. The degree of glucosylation at each position is close to stoichiometric, but the first O-unit of the O-polysaccharide chain linked to the LPS core lacks any glucosyl residue [39-41].

In serotype 7, GlcNAc carries the α-D-Glc*p*-(1→2)-α-D-Glc*p*-(1→ disaccharide [37, 38], which defines O-factor VII (IC). Subtype 7a was originally called 1c [37], but wild-type strains of this subtype do not react with antibodies against O-factor I. Therefore, it was suggested to rename it 7a and to replace the type O-factor IC with VII in the serotyping scheme of *S. flexneri* [38].

O-Acetylation has been identified on Rha^I, Rha^{III}, and GlcNAc [10, 22-24, 33, 38, 42] (Fig. 1). The degree of O-acetylation is variable and depends likely not only on strain but also on storage and cultivation conditions. On Rha^I, the 2-O-acetylation is stoichiometric or close to stoichiometric, whereas on GlcNAc, 6-O-acetylation varies from 30 to 75%. Rha^{III} is O-acetylated at position 3 in some O-units and at position 4 in some others (3/4-O-acetylation), the former being the major (25-70%) and the latter the minor (15-25%) modification site. All combinations of O-acetylated and non-acetylated Rha^{III} and GlcNAc have been found in the O-units of the serotype 2a O-polysaccharide, and hence O-acetylation on both residues is random [20]. In a short-chain LPS having a single O-unit, GlcNAc lacks 6-O-acetylation and Rha^{III} is mono-O-acetylated at any position [40].

O-Acetylation on Rha^I, Rha^{III}, and GlcNAc defines group O-factors 6, 9, and 10, respectively. Serotype 3 strains express type O-factor III, which depends on the same 2-O-acetyl group on Rha^I as group O-factor 6 [42] but, in contrast to the latter, is abolished by glucosylation on GlcNAc in serotypes 1b, 4b, and 7b. O-Factor III could be recovered from serotypes 1b and 4b by transformation with the functional *oacD* gene, which resulted in partial 6-O-acetylation (~25-30%) accompanied by deglucosylation on GlcNAc [24]. 6-O-Acetylation on GlcNAc also occurs in the common enterobacterial polysaccharide antigen [43, 44] and, as a result, O-factor 10 is expressed by some other enteric bacteria, including *Shigella sonnei* phase II [24].

In early studies of *S. flexneri* O-antigen structures, O-acetylation on Rha^{III} and GlcNAc has been overlooked, and the corresponding O-factors 9 and 10 have not been included in the *S. flexneri* serotyping scheme. To fill the gap, we suggest to further divide the existent serotypes into O-factor 9- and 10-positive and -negative subtypes (i) by keeping the old names for the subtypes that lack both O-factors 9 and 10 [24, 33] and for serotype 1b whose O-factor 9-negative variant has not been found in nature, and (ii) by indication of expression of one or both of the O-factors 9 and 10, by adding subscript 1 or 2, respectively (e.g. $2a_1$ and $2a_2$ for subtypes characterized by the antigenic formulae II: 10 and II: 9; 10, respectively). To distinguish these subtypes, monospecific antisera against O-factors 9 [33] and 10 [24] that have already been generated and verified should be included into the serological diagnostic kit.

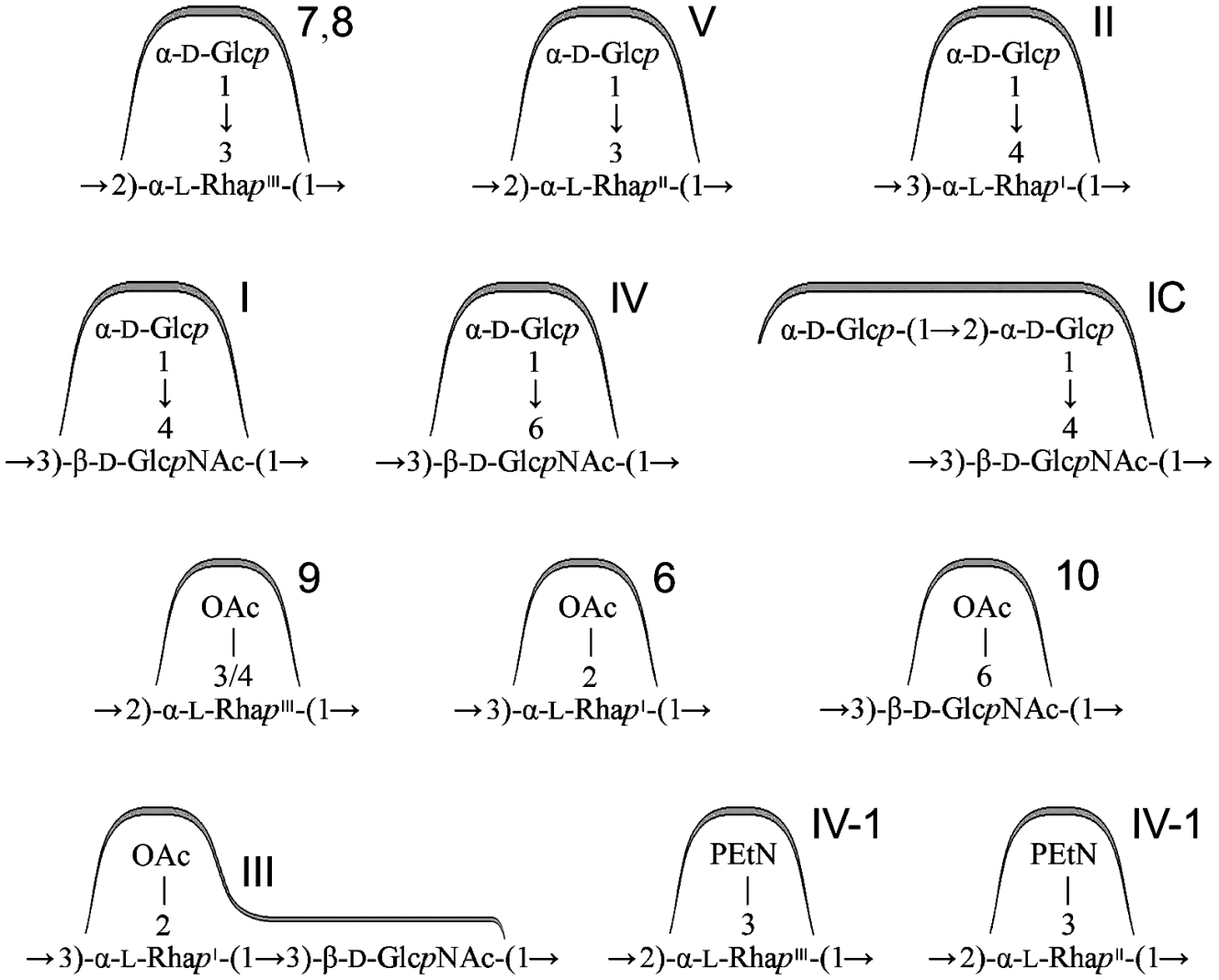


Fig. 1. Immunodeterminants (O-factors) associated with various O-antigen-modifying groups.

Phosphorylation with PEtN has been reported in subtypes 4av, Xv, Yv, and Yv_1 designated as "variant" subtypes by adding letter "v" to the names of the corresponding PEtN-positive subtypes [26-29]. Stoichiometric phosphorylation occurs at position 3 of Rha^{III} in subtype 4av [26] or Rha^{II} in subtype Xv [27] (Fig. 1) with minor phosphorylation (~10%) on the neighboring rhamnose residue, in subtype Xv the minor PEtN group replacing the glucosyl group on Rha^{III} [27]. In subtypes Yv and Yv_1, both rhamnose residues are phosphorylated, one being completely (Rha^{II} in subtype Yv or Rha^{III} in subtype Yv_1) and the other partially modified ([28] and authors' unpublished data). It has been demonstrated that in subtype Yv_1, bisphosphorylation occurs in the non-O-acetylated O-units only, and the Yv_1 O-polysaccharide is composed of blocks of repeats differing in the number of PEtN groups and the presence or absence of O-acetylation [28].

O-Acetylation on Rha^{III} and GlcNAc as well as phosphorylation with PEtN has not been identified in early structural studies of *S. flexneri* O-antigens, but monoclonal antibody MASF IV-1 generated against a PEtN-positive strain of subtype 4a (now 4av) has been included into the MASF reagents (Reagensia AB). This antibody is useful for detection of all PEtN-positive strains, as the corresponding group O-factor IV-1 (originally called 4X [31]) is defined by phosphorylation with PEtN whether it occurs on Rha^{II} or Rha^{III} [27, 29]. That the same monoclonal antibody recognizes a PEtN-associated epitope on either of the two monosaccharides is probably due to a sharp turn of the polysaccharide chain at each 2-substituted Rha residue, which makes both Rha^{II} and Rha^{III} easily accessible to interaction with the protein and neglects the role of the neighboring sugar residues.

The serotype 6 O-polysaccharide (2) is acidic due to the presence of D-galacturonic acid (GalA) [10, 36]. The

first monosaccharide in the serotype 6 O-unit is 2-acetamido-2-deoxy-D-galactose (GalNAc) rather than GlcNAc, but the α1→2-linked rhamnose disaccharide at the other side of the O-unit is shared by all *S. flexneri* serotypes.

$$\rightarrow 2)\text{-}\alpha\text{-L-Rha}p^{III}\text{-}(1\rightarrow 2)\text{-}\alpha\text{-L-Rha}p^{II}\text{-}(1\rightarrow 4)\text{-}$$

$$\text{-}\beta\text{-D-Gal}p\text{A-}(1\rightarrow 3)\text{-}\beta\text{-D-Gal}p\text{NAc-}(1\rightarrow \qquad (2)$$

Serotype 6 strains are recognized by typing antiserum VI specific to an unidentified O-polysaccharide domain. The only modification of the serotype 6 O-antigen that has been characterized chemically is 3/4-O-acetylation on Rha^{III} [10, 33]. Due to the presence of an O-polysaccharide backbone fragment in common with non-6 serotypes of *S. flexneri*, this modification confers serotype 6 with O-factor 9, which is recognized smoothly by antiserum 9 produced against an O-factor 9-positive strain of serotype 2 [23, 33]. O-Factor 9 is present in all serotype 6 strains tested [10, 23, 33]. As in serotype 2a, the terminal Rha^{III} residue of the single O-unit in a short-chain serotype 6 LPS is O-acetylated randomly [40].

Seven atypical serotype 6 strains collected in Bangladesh during 1985-1987 [45] and 1997-2000 [46] were recognized by monoclonal antibody MASF IV-1, suggesting that subtype 6v carrying a PEtN-associated epitope emerged in nature. Structural and genetic bases of the O-factor IV-1 expression in this subtype remain to be elucidated.

GENETIC BASIS OF O-ANTIGEN MODIFICATIONS AND RESULTANT SEROTYPE CONVERSION

Glucosylation. Three Gtr proteins (GtrA, GtrB, and type-specific Gtr (Gtr(type)) mediate glucosylation of the O-polysaccharide backbone (1). GtrA and GtrB are highly conserved and functionally interchangeable between serotypes. GtrB catalyzes synthesis of undecaprenyl phosphate-β-glucose (UndP-β-Glc) from UDP-α-Glc, and GtrA functions as flippase allowing translocation of the UndP-β-Glc from cytoplasm to periplasm. The third protein, Gtr(type), is a serotype-specific glucosyltransferase (GtrI, GtrII, GtrIV, GtrV, GtrVII (formerly GtrIc), and GtrX) responsible for the transfer of the glucosyl group from UndP-β-Glc to a certain position of one of the sugar residues of the growing O-polysaccharide chain. The Gtr(type) enzymes are integral membrane proteins consisting of 8-10 transmembrane helices with the active sites located in the large periplasmic loops at the N- and C-termini. They have weak similarity to other known glucosyltransferase families and are predicted to be members of the GT-C superfamily, which utilize a phospholipid-activated donor sugar substrate [47].

A single operon on the chromosome encoding Gtr proteins (*gtr* cluster) is carried by a (cryptic) prophage acquired by lysogeny of the bacteria with one or two from five temperate bacteriophages (SfI, SfII, SfIV, SfV, and SfX) [14, 48]. Each prophage (or a couple of prophages in the case of bisglucosylation in the O-unit) is integrated into the *thrW* tRNA gene in the same region adjacent to the *proA* gene in the *S. flexneri* genome. The *gtr* cluster is localized immediately downstream of the phage attachment site *attP*, which follows the integrase (*int*) and excisionase (*xis*) genes. All bacteriophages have been isolated from the corresponding *S. flexneri* strains and well characterized [41, 46-54].

Lysogeny with bacteriophages SfI, SfII, SfIV, SfV, and SfX converts serotype Y to serotypes 1a, 2a, 4a, 5a, and X, respectively (Fig. 2), whereas the potential recipient range among other serotypes is quite different. This is single serotype X for SfI, two serotypes (3b and 5a) for SfII, two serotypes ($2a_1$ and 3b) for SfX, four subtypes of serotypes 1a, VII, and X for SfIV, and six subtypes of serotypes 1-4 for SfV of the 12 serotypes tested. The limitation in the host recognition is evidently due to the phage immunity from a modified O-antigen, which constitutes the receptor for the phage adsorption on the cell surface, a mechanism by which lysogeny prevents subsequent infection of bacteria by homologous or related phages ([52-54], Q. Sun, J. Wang, X. Luo et al., unpublished data). Accordingly, the order of lysogeny with two phages giving rise to serotypes carrying more than one phage-borne modification factor can be constrained; for instance, phage SfI can infect SfX-carrying serotype X strains giving rise to serotype 1d, but serotype 1 strains are resistant to phage SfX [55] (Fig. 2).

Inactivating mutations in the *gtr* locus occur in a number of wild-type strains that carry a serotype-converting phage, resulting in their reversion to the parental serotype (Y) or (in case of bisglucosylation) an intermediate serotype (Fig. 2). For instance, from 35 serotype Y strains, 13 strains possess defective gene *gtrII* and six strains defective gene *gtrI*. From 19 strains of O-factor IV-1-positive serotype Yv and Yv_1, 13 strains have mutations in either one or both genes *gtrII* and *gtrB*, and three strains possess a defective *gtrX* [29]. As a result, the same serotype may have multiple origins, e.g. subtypes Yv and Yv_1 emerged independently at least three times from serotypes Y, Xv, and 2a by acquisition of an *opt*-carrying plasmid, inactivation of a *gtr* gene, and both events, respectively [29] (Fig. 2).

In serotype 7 that is distinguished by the presence of the side-chain α1→2-linked glucose disaccharide, the addition of the first glucosyl group is mediated by the same *gtr* cluster within a SfI prophage as in serotype 1 [56]. The *gtrVII* gene designated originally *gtrIC* codes for the serotype 7-specific glucosyltransferase that mediates addition of the second glucose residue to the first one. As

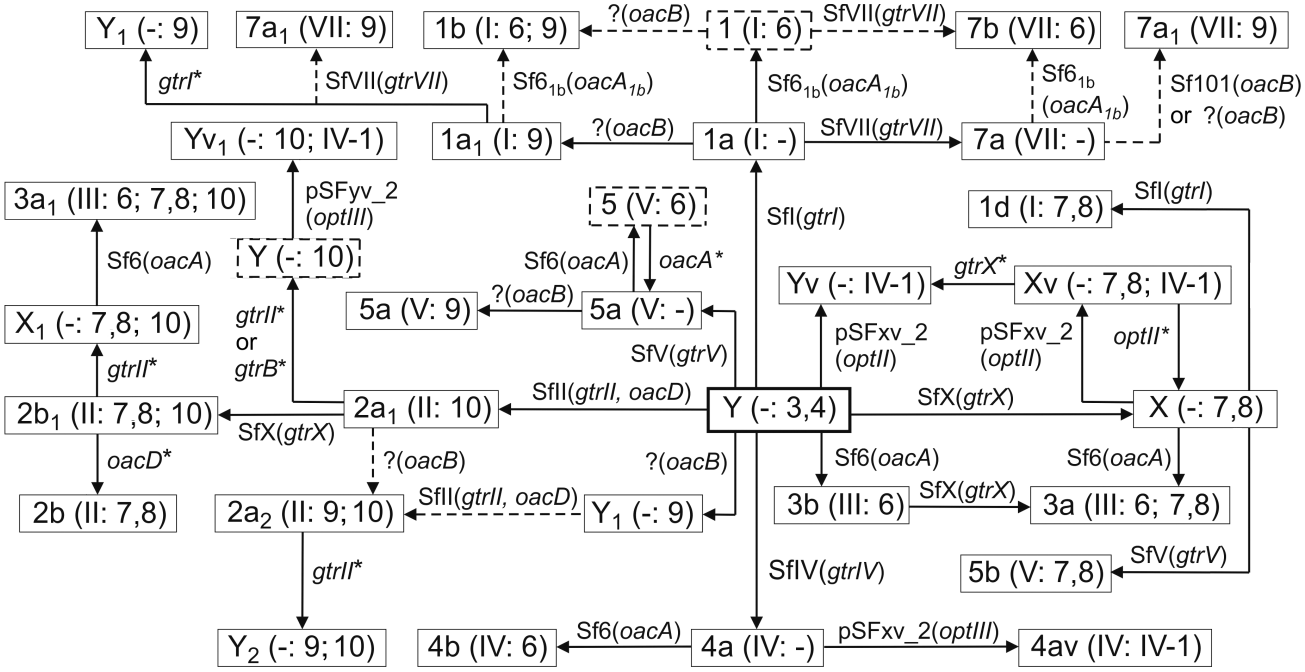


Fig. 2. Conversion pathways of serotypes 1-5, 7, X, and Y mediated by bacteriophages SfI, SfII, SfIV, SfV, SfVII, SfX, Sf6 or $Sf6_{1b}$, Sf101, and plasmids pSFxv_2 or pSFyv_2. Adopted from ([24, 25, 27, 29, 52-56, 59], Q. Sun, J. Wang, X. Luo et al., unpublished data). Antigenic formulae are shown in parentheses (except for serotype Y, group O-factor 3,4 associated with the O-polysaccharide backbone is omitted). Asterisk indicates gene inactivation. Dashed arrows show that the order of serotype-converting events is unknown. Putative intermediates that have not been found in nature are shown in dashed rectangular box. Hypothetical bacteriophages $Sf6_{1b}$ and SfVII have not been isolated. The bacteriophage Sf101 origin of the *oacB* gene has been demonstrated for two serotype $7a_1$ strains, and mobilization of *oacB* into other 3/4-O-acetylation-carrying 7a and non-7a strains suggested to result from disruption of the Sf101 prophage by IS (insertion sequence) elements followed by recombination [25].

the other type-specific *gtr(type)* genes, *gtrVII* is present as part of a three-gene *gtr* cluster but is located at a different place on the chromosome adjacent to the conserved *yejO* locus. It is distantly related to the other *S. flexneri gtr* clusters and appears to have been acquired from outside the species, presumably via infection by a hypothetical bacteriophage SfVII (SfIC) [56].

O-Acetylation. 2-O-Acetylation of Rha^I is mediated by an acetyltransferase, which was originally named Oac, but after discovery of other O-antigen-modifying acetyltransferases in *S. flexneri*, it was suggested to rename it OacA [22]. The receptor for OacA is the O-antigen of serotype Y having the basal structure (1) as well as some other serotypes (Fig. 2). OacA consists of 10 α-helical membrane-spanning regions with both the N- and C-termini located in the cytoplasm. It bears homology to several known and predicted acetyltransferases with most homology existing in the N-terminal transmembrane regions [57]. In serotypes 3a, 3b, and 4b, the *oacA* gene and an adjacent integrase-encoding gene are carried by the temperate bacteriophage Sf6, which, like the *gtr* locus-carrying bacteriophages, is a member of the canonical lambdoid phage group. The Sf6 genome is integrated into the *argW* tRNA gene of the host chromosome next to the conserved *yfdC* gene (Fig. 3a) [58].

Serotypes 1b [59] and 7b (authors' unpublished data) possess a variant *oacA* gene named $oacA_{1b}$ (originally oac_{1b}), which shares with *oacA* 88-89% identity at the DNA level and 85% identity at the protein level. Despite the rather high sequence variation, *oacA* and $oacA_{1b}$ are functionally interchangeable in 2-O-acetylation of Rha^I. $oacA_{1b}$ is located in a chromosomal region between the conserved *torT* and *ycmA* genes, which evidently has a phage origin but is different from the Sf6 phage genome (Fig. 3a). Whereas serotypes 3a, 3b, and 4b can be generated by infecting with bacteriophage Sf6 strains of serotypes X, Y, and 4a, respectively, this phage cannot convert serotype 1a into serotype 1b. Therefore, it is likely that $oacA_{1b}$ has been obtained from outside *S. flexneri*, probably by infection with another bacteriophage (hypothetical $Sf6_{1b}$) rather than evolved by divergence from the *oacA* gene.

6-O-Acetylation of GlcNAc is mediated by an *oac* homolog designated *oacD*, which is carried by SfII bacteriophage also responsible for 4-glucosylation of Rha^I giving rise to serotype 2 (Fig. 3b) [24]. The occurrence of an insertion sequence (IS) upstream of *oacD* suggests that this gene was incorporated into the SfII genome by an insertion event. The functional *oacD* gene also is present in strains of several non-2 serotypes ($3a_1$, X_1, Y_1, Y_2, and

Yv_1) that carry a cryptic SfII prophage with a dysfunctional *gtr* locus for the type II glucosylation. Therefore, the OacD-mediated 6-O-acetylation of GlcNAc does not depend on the functional *gtr* locus.

3/4-O-Acetylation of Rha^{III} that occurs in subtypes $1a_1$, 1b, $2a_2$, $5a_1$, 7_1, Y_1, 6, and Y_2 [22, 25] is mediated by another Oac homolog called OacB. This modification was unaffected by transformation of $2a_2$ and Y_1 strains with either the *gtrABX* locus for 3-O-glucosylation (authors' unpublished data) or the *optII* gene for 3-O-phosphorylation [60] of Rha^{III}. In contrast, transformation of 2b and X strains with *oacB* from a $2a_2$ strain resulted in their conversion into serotypes 2a and Y, respectively, due to replacement of 3-O-glucosylation with 3/4-O-acetylation on Rha^{III} [22]. The mechanism that makes the O-acetylation the preferable modification on Rha^{III} remains to be elucidated.

Two alternative locations of the *oacB* gene on the bacterial chromosome have been reported. In some subtype $7a_1$ strains, *oacB* is carried by Sf101 prophage integrated in the *sbcB* gene next to *yeeD* [22] (Fig. 3b). In several different subtype $7a_1$ strains [25] and 3/4-O-acetylation-carrying strains of other serotypes [22], *oacB* maps upstream of the *adrA* gene in the same *proA-adrA* region on the chromosome, in which the *gtr*-carrying prophages are integrated. It is located downstream of an integrase-encoding gene (*int*), and the *int-oacB* locus is flanked by IS elements giving rise to a transposon-like structure [22]. In serotype 2a strains examined, this structure is located immediately downstream of the SfII prophage genome (Fig. 3b).

In serotype 6, yet another *oac* homolog, *oacC*, common for all strains of this serotype, is responsible for the 3/4-O-acetylation of Rha^{III} (Fig. 3c) [23]. It maps in a

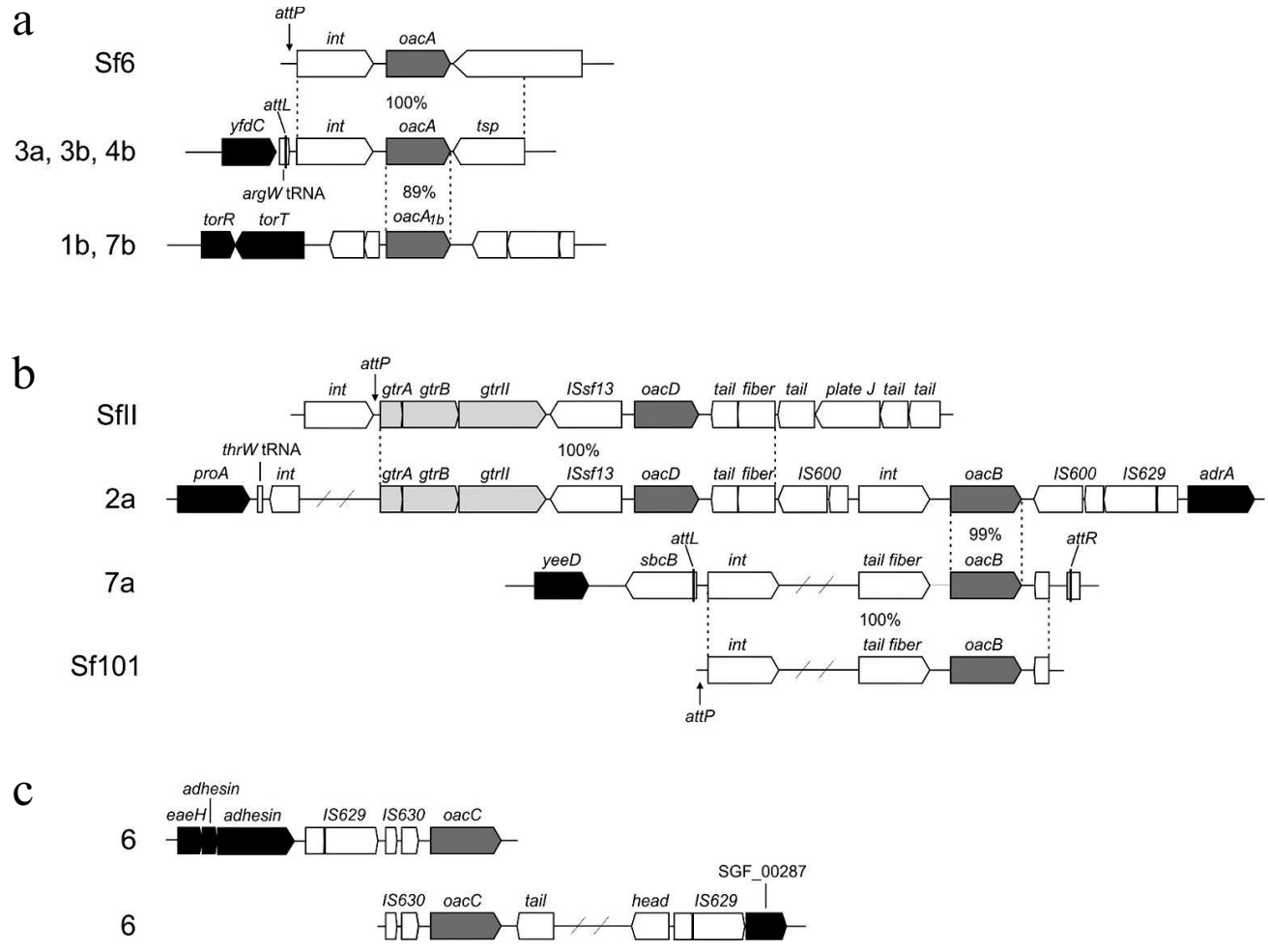


Fig. 3. Organizations of the genomic regions of bacteriophages and *S. flexneri* carrying the O-acetyltransferase genes: *oacA* (former *oac*) in bacteriophage Sf6 and serotypes 3a (subtype $3a_1$), 3b, 4b, and $oacA_{1b}$ (former oac_{1b}) in serotypes 1b and 7b (a); *oacD* in bacteriophage SfII and serotype 2a (subtype $2a_2$, strain Sf301), *oacB* in bacteriophage Sf101, serotype 7a (subtype $7a_1$, strain SFL1683) and 2a (b); *oacC* in serotype 6 (strains CCH 060 (top) and CDC 796-83 (bottom)) (c). Adopted from [22-25, 57, 59]. The *gtr* locus for type II glucosylation, the O-acetyltransferase genes, and the conserved flanking genes are shown in light gray, dark gray, and black, respectively. *attP* indicates attachment site on phage, *attL* and *attR* indicate left and right ends of the integrated phage genome.

phage-like structure localized in yet another place on the chromosome (Fig. 3c). OacB and OacC have high sequence homology (72% identity) and interchangeable function in mediating the 3/4-O-acetylation of Rha^{III}. This is not surprising as the O-polysaccharides of all *S. flexneri* serotypes share a →2)-α-L-Rha*p*^{III}-(1→2)-α-L-Rha*p*^{II}-(1→ disaccharide fragment, which evidently serves as the acceptor substrate for both OacB and OacC. The three known rhamnose-modifying acyltransferases OacA-OacC present higher homology in the regions conserved among the inner membrane *trans*-acylase family proteins [23]; particularly, conserved are amino acid residues R73 and R76, which are known to be critical for Oac functioning [57]. The divergent *oac* genes might have been gained from different bacterial species in independent events.

Phosphorylation with PEtN. A polymorphic *opt* gene (originally called *lpt-O*) encoding for PEtN transferases is responsible for adding a PEtN group to Rha^{II} or/and Rha^{III} in subtypes 4av, Xv, Yv, and Yv_1 [27, 28]. The Opt proteins have been predicted to belong to the sulfatase superfamily and to contain a sulfatase domain on the carboxyl terminus, which putatively is involved in catalyzing the transfer of PEtN to a sugar residue. Two functionally interchangeable *opt* alleles, *optII* and *optIII*, are borne by double-stranded circular plasmids 6850 bp in length called pSFxv_2 and pSFyv_2, respectively (Fig. 4). OptII and OptIII preferentially mediate phosphorylation of Rha^{II} and Rha^{III} [27-29], and, accordingly, *optII* and *optIII* are present in subtypes Xv and 4av, respectively. An explanation for the evolution of this gene may be selection pressure as in serotype 4a, Rha^{III} is not occupied, and the OptIII can easily mediate the addition of PEtN onto it. In contrast, in serotype X, Rha^{III} carries a glucosyl group, and the OptIII-mediated phosphorylation may not effectively compete for Rha^{III} with type 7,8 glucosylation, whereas OptII can smoothly modify Rha^{II}. In subtypes Yv and Yv_1, the *opt* form depends on the strain origin: this is *optII* in Yv strains derived from serotype Y or Xv or *optIII* in serotype 2-derived Yv_1 strains (Fig. 2) [29]. Accordingly, Rha^{II} and Rha^{III} are predominantly phosphorylated in subtypes Yv and Yv_1, respectively ([28] and authors' unpublished data).

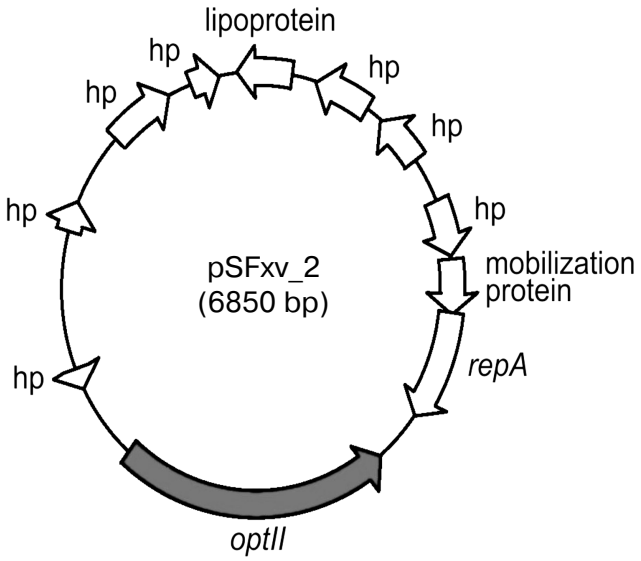


Fig. 4. Genomic structure of plasmid pSFxv_2 carrying the PEtN transferase gene *optII*. Adopted from [27]. *repA*, replication initiation protein gene; hp, hypothetical protein.

It has been demonstrated that plasmids pSFxv_2 [60] and pSFyv_2 (authors' unpublished data) can be transferred into, and stably maintained in, strains of other *S. flexneri* serotypes (1 to 6) giving rise to unnatural O-factor IV-1-positive serovariants. Upon the transformation, the initial serospecificity is either retained or lost, or manifestation of an initial epitope(s) may be weakened. Phosphorylation with PEtN may interfere with other modifications on the O-antigen not only on the same monosaccharide but also on different sugar residues; for instance, 3-O-phosphorylation on Rha^{II} is incompatible with 4-O-glucosylation on Rha^{I} [60]. This phenomenon seems to account for the absence of the PEtN-carrying variant of 2a (subtype 2av) in nature, while its non-phosphorylated form is highly prevalent among clinical isolates. It also shows that in the emergence of subtype Yv_1 from subtype $2a_1$, inactivation of *gtrII* or *gtrB* and the resultant loss of type II glucosylation occurred first and pSFyv_2 plasmid was gained subsequently (Fig. 2).

DISCUSSION

Known modifications of the O-antigens of *S. flexneri* involve glucosylation, O-acetylation, or/and phosphorylation with PEtN at various monosaccharides in the O-unit. None is unique on its own, but their most diverse combinations provide an unusually high diversity of the O-antigen forms. By now, about 30 structural variants of the O-antigens of *S. flexneri* with the same basal structure have been identified, which exceed significantly their number in any other bacteria.

Glucosylation has been known to take place at the periplasmic side of the inner membrane together with the O-unit polymerization [14]. The data on the O-acetylation summarized earlier [14] are contradictory, and it remains unknown where phosphorylation of the O-antigen takes place. Some modifications on different monosaccharide residues, e.g. 4-O-glucosylation on Rha^{I} and 3-O-phosphorylation on Rha^{II}, are incompatible with each other. The modification pattern of the first O-unit that is linked to the LPS core may be different; particularly, glucosylation only happens to at least the second O-unit from the growing end, the first one being unaltered [40]. This finding and the fact that the same Wzy polymerase catalyzes synthesis of structurally different O-

polysaccharides indicate complex interactions of the enzymes involved in the O-antigen polymerization and modifications, which remain to be elucidated.

The diverse O-polysaccharide modifications are encoded by a number of genes outside the O-antigen clusters between *galF* and *gnd*, and their roles have been elucidated by serological analysis of the wild-type and non-polar mutants and structure determination of their isolated O-polysaccharides. Multiple genetic mechanisms have been recognized to underlie the *S. flexneri* O-antigen modifications and the resultant serotype conversions.

Most common is lysogeny by bacteriophages that encode glucosyltransferases and/or acetyltransferases. One of the bacteriophages, SfII, includes both *gtr* gene cluster for 4-O-glucosylation of Rha^{I} and *oacD* gene for 6-O-acetylation of GlcNAc [24]. This is a unique situation in *S. flexneri* that one serotype-converting phage carries two genetic factors involved in different types of O-antigen modifications.

Another way of O-antigen modification factor mobilization is a combination of the *oacB* gene for 3/4-O-acetylation of Rha^{III} with several IS elements giving rise to a transposon-like structure [22], which could evolve as a result of disruption of *oacB*-carrying Sf101 prophage by IS elements [25].

Most *gtr*-containing prophages and the *oacB*-carrying locus map in the same region on the chromosome upstream of the *adrA* gene, which seems to be a conserved insertion site for mobile genetic elements. On the other hand, the *gtr* locus for addition of the second glucose residue in serotype 7 as well as *oacB* in some subtype $7a_1$ strains and genes for all other acetyltransferases are parts of bacteriophage genomes or phage-like structures integrated at different places on the chromosome. The polymorphic *opt* gene that encodes for PEtN-transferases responsible for phosphorylation of Rha^{III} or/and Rha^{II}, is carried on 6.85-kb plasmids, which have high dissemination potential among *S. flexneri* serotypes ([60] and authors' unpublished data).

Therefore, the primary genetic mechanism of diversification of *S. flexneri* non-serotype 6 O-antigen structures is acquisition of various transferable genetic factors, including prophages and plasmids, which can easily spread among different strains. Inactivation of a gene involved in an O-antigen modification, such as a *gtr* gene, *oac*, or *opt*, contributes to further conversion of serotypes. Because of the involvement of multiple O-antigen modification factors, the same serotype may emerge multiple times not only by the same [61] but also by different ways, including both gaining and losing one or several factors (Fig. 2).

Recent identification of PEtN phosphorylation and O-acetylation at new sites has refined the notion of the antigenic heterogeneity of *S. flexneri* (table and Fig. 2). It has become clear that the diversity of the O-antigenic forms in these bacteria was underestimated in the past. For their separate detection, it was proposed to apply specific antisera obtained by absorption of immune sera against wild-type strains with the corresponding isogenic mutants or vice versa [23, 24, 27, 28, 33, 62, 63], as well as molecular approaches using as targets the specific genes responsible for the O-antigen modifications [22-24, 27, 28, 61]. Serological and molecular screening showed that most newly found O-antigen forms occur rather frequently. The data show the expediency of an extension of the existing *S. flexneri* serotyping scheme by recognition of the representative strains of the new variants as distinctive subtypes.

Human immune response to *S. flexneri* infection is serotype specific with protection against subsequent infection by the same serotype only. As acquisition of multiple drug resistance, the appearance of new surface epitopes due to O-antigen modifications would be expected to offer a significant advantage to the pathogen and to promote its spread in human populations. For instance, serotype Xv distinguished by PEtN phosphorylation of Rha^{II} appeared initially in one province in China in 2001 and rapidly expanded to most provinces, surpassing 2a as the predominant serotype [62]. It has been demonstrated that glucosylation on Rha^{I}, Rha^{II}, and GlcNAc confers a specific advantage on *S. flexneri* [16], and it may be suggested that 3/4-O-acetylation on Rha^{III} is somehow beneficial to the bacteria too, as it occurs in >95% strains of serotypes 1a, 1b, and 2a, which are predominant in developing countries.

Therefore, the data presented in this review, including elucidation of the finer details of the O-antigen modifications and the underlying genetic mechanisms, shed light, and provide avenues for further studies, on the role of the O-antigen variations in the antigenicity, pathogenicity, and epidemicity of *S. flexneri*. They also have profound implications in development of improved diagnostic methods and efficient shigellosis vaccines targeting newly discovered genes and epitopes.

We acknowledge our colleagues who have significantly contributed to the advancement of this field.

This work was supported by the Russian Science Foundation (Grant No. 14-14-01042 for Y. A. K.). Q. S. and J. X. were supported by the National Natural Science Foundation of China (No. 81271788, 81290340 and 81290345) and the National Key Program for Infectious Diseases of China (2013ZX10004221, 2013ZX10004216-001-002).

REFERENCES

1. Bardhan, P., Faruque, A. S., Naheed, A., and Sack, D. A. (2010) Decrease in shigellosis-related deaths without *Shigella* spp.-specific interventions, Asia, *Emerg. Infect. Dis.*, **16**, 1718-1723.

2. Schroeder, G. N., and Hilbi, H. (2008) Molecular pathogenesis of *Shigella* spp.: controlling host cell signaling, invasion, and death by type III secretion, *Clin. Microbiol. Rev.*, **21**, 134-156.
3. Ewing, W. H., and Lindberg, A. A. (1984) Serology of *Shigella*, *Methods Microbiol.*, **14**, 113-142.
4. Pupo, G. M., Lan, R., and Reeves, P. R. (2000) Multiple independent origins of *Shigella* clones of *Escherichia coli* and convergent evolution of many of their characteristics, *Proc. Natl. Acad. Sci. USA*, **97**, 10567-10572.
5. Kotloff, K. L., Winickoff, J. P., Ivanoff, B., Clemens, J. D., Swerdlow, D. L., Sansonetti, P. J., Adak, G. K., and Levine, M. M. (1999) Global burden of *Shigella* infections: implications for vaccine development and implementation of control strategies, *Bull. World Health Organ.*, **77**, 651-666.
6. Shiferaw, B., Shallow, S., Marcus, R., Segler, S., Soderlund, D., Hardnett, F. P., and Van Gilder, T. (2004) Trends in population-based active surveillance for shigellosis and demographic variability in FoodNet sites, 1996-1999, *Clin. Infect. Dis.*, **38** (Suppl. 3), S175-S180.
7. Livio, S., Strockbine, N., Panchalingam, S., Tennant, S. M., Barry, E. M., Marohn, M. E., Antonio, M., Hossain, A., Mandomando, I., Ochieng, J., B. Oundo, J. O., Qureshi, S., Ramamurthy, T., Tamboura, B., Adegbola, R. A., Hossain, M. J., Saha, D., Sen, S., Faruque, A. S., Alonso, P. L., Breiman, R. F., Zaidi, A. K., Sur, D., Sow, S. O., Berkeley, L. Y., O'Reilly, C., Mintz, E. D., Biswas, K., Cohen, D., Farag, T. H., Nasrin, D., Wu, Y., Blackwelder, W. C., Kotloff, K. L., Nataro, J. P., and Levine, M. M. (2014) *Shigella* isolates from the Global Enteric Multicenter Study inform vaccine development, *Clin. Infect. Dis.*, **59**, 933-941.
8. Kenne, L., Lindberg, B., Petersson, K., Katzenellenbogen, E., and Romanowska, E. (1978) Structural studies of *Shigella flexneri* O-antigens, *Eur. J. Biochem.*, **91**, 279-284.
9. Simmons, D. A. R., and Romanowska, E. (1987) Structure and biology of *Shigella flexneri* O antigens, *J. Med. Microbiol.*, **23**, 289-302.
10. Perepelov, A. V., Shekht, M. E., Liu, B., Shevelev, S. D., Ledov, V. A., L'vov, V. L., Senchenkova, S. N., Shashkov, A. S., Feng, L., Aparin, P. G., Wang, L., and Knirel, Y. A. (2012) *Shigella flexneri* O-antigens revisited: final elucidation of the O-acetylation profiles and a survey of the O-antigen structure diversity, *FEMS Immunol. Med. Microbiol.*, **66**, 201-210.
11. Liu, B., Knirel, Y. A., Feng, L., Perepelov, A. V., Senchenkova, S. N., Wang, Q., Reeves, P. R., and Wang, L. (2008) Structure and genetics of *Shigella* O antigens, *FEMS Microbiol. Rev.*, **32**, 627-653.
12. Li, Y., Cao, B., Liu, B., Liu, D., Gao, Q., Peng, X., Wu, J., Bastin, D. A., Feng, L., and Wang, L. (2009) Molecular detection of all 34 distinct O-antigen forms of *Shigella*, *J. Med. Microbiol.*, **58**, 69-81.
13. Perepelov, A. V., Shevelev, S. D., Liu, B., Senchenkova, S. N., Shashkov, A. S., Feng, L., Knirel, Y. A., and Wang, L. (2010) Structures of the O-antigens of *Escherichia coli* O13, O129 and O135 related to the O-antigens of *Shigella flexneri*, *Carbohydr. Res.*, **345**, 1594-1599.
14. Allison, G. E., and Verma, N. K. (2000) Serotype-converting bacteriophages and O-antigen modification in *Shigella flexneri*, *Trends Microbiol.*, **8**, 17-23.
15. Sun, Q., Lan, R., Wang, Y., Zhao, A., Zhang, S., Wang, J., Xia, S., Jin, D., Cui, Z., Zhao, H., Li, Z., Ye, C., Jing, H., and Xu, J. (2011) Development of a multiplex PCR assay targeting O-antigen modification genes for molecular serotyping of *Shigella flexneri*, *J. Clin. Microbiol.*, **49**, 3766-3770.
16. West, N. P., Sansonetti, P., Mounier, J., Exley, R. M., Parsot, C., Guadagnini, S., Prevost, M. C., Prochnicka-Chalufour, A., Delepierre, M., Tanguy, M., and Tang, C. M. (2005) Optimization of virulence functions through glucosylation of *Shigella* LPS, *Science*, **307**, 1313-1317.
17. Kohler, H., Rodrigues, S. P., and McCormick, B. A. (2002) *Shigella flexneri* interactions with the basolateral membrane domain of polarized model intestinal epithelium: role of lipopolysaccharide in cell invasion and in activation of the mitogen-activated protein kinase ERK, *Infect. Immun.*, **70**, 1150-1158.
18. Hong, M., and Payne, S. M. (1997) Effect of mutations in *Shigella flexneri* chromosomal and plasmid-encoded lipopolysaccharide genes on invasion and serum resistance, *Mol. Microbiol.*, **24**, 779-791.
19. Brahmbhatt, H. N., Lindberg, A. A., and Timmis, K. N. (1992) *Shigella* lipopolysaccharide: structure, genetics, and vaccine development, *Curr. Top. Microbiol. Immunol.*, **180**, 45-64.
20. Kubler-Kielb, J., Vinogradov, E., Chu, C., and Schneerson, R. (2007) O-Acetylation in the O-specific polysaccharide isolated from *Shigella flexneri* serotype 2a, *Carbohydr. Res.*, **342**, 643-647.
21. Perepelov, A. V., L'vov, V. L., Liu, B., Senchenkova, S. N., Shekht, M. E., Shashkov, A. S., Feng, L., Aparin, P. G., Wang, L., and Knirel, Y. A. (2009) A similarity in the O-acetylation pattern of the O-antigens of *Shigella flexneri* types 1a, 1b and 2a, *Carbohydr. Res.*, **344**, 687-692.
22. Wang, J., Knirel, Y. A., Lan, R., Senchenkova, S. N., Luo, X., Perepelov, A. V., Wang, Y., Shashkov, A. S., Xu, J., and Sun, Q. (2014) Identification of an O-acyltransferase gene (*oacB*) that mediates 3- and 4-O-acetylation of rhamnose III in *Shigella flexneri* O antigens, *J. Bacteriol.*, **196**, 1525-1531.
23. Knirel, Y. A., Luo, X., Wang, J., Senchenkova, S. N., Lan, R., Shpirt, A. M., Du, P., Shashkov, A. S., Xu, J., and Sun, Q. (2014) Genetic and structural identification of an O-acyltransferase gene (*oacC*) responsible for the 3/4-O-acetylation on rhamnose III in *Shigella flexneri* serotype 6, *BMC Microbiol.*, **14**, 266.
24. Sun, Q., Knirel, Y. A., Wang, J., Luo, X., Senchenkova, S. N., Lan, R., Shashkov, A. S., and Xu, J. (2014) Serotype-converting bacteriophage SfII encodes an acyltransferase protein that mediates 6-O-acetylation of GlcNAc in *Shigella flexneri* O-antigens, conferring on the host a novel O-antigen epitope, *J. Bacteriol.*, **196**, 3656-3666.
25. Jakhetia, R., Marri, A., Stahle, J., Widmalm, G., and Verma, N. K. (2014) Serotype-conversion in *Shigella flexneri*: identification of a novel bacteriophage, Sf101, from a serotype 7a strain, *BMC Genomics*, **15**, 742.
26. Perepelov, A. V., L'vov, V. L., Liu, B., Senchenkova, S. N., Shekht, M. E., Shashkov, A. S., Feng, L., Aparin, P. G., Wang, L., and Knirel, Y. A. (2009) A new ethanolamine phosphate-containing variant of the O-antigen of *Shigella flexneri* type 4a, *Carbohydr. Res.*, **344**, 1588-1591.
27. Sun, Q., Knirel, Y. A., Lan, R., Wang, J., Senchenkova, S. N., Jin, D., Shashkov, A. S., Xia, S., Perepelov, A. V., Chen,

Q., Wang, Y., Wang, H., and Xu, J. (2012) A novel plasmid-encoded serotype conversion mechanism through addition of phosphoethanolamine to the O-antigen of *Shigella flexneri*, *PLoS ONE*, **7**, e46095.
28. Knirel, Y. A., Lan, R., Senchenkova, S. N., Wang, J., Shashkov, A. S., Wang, Y., Perepelov, A. V., Xiong, Y., Xu, J., and Sun, Q. (2013) O-antigen structure of *Shigella flexneri* serotype Yv and effect of the *lpt-O* gene variation on phosphoethanolamine modification of *S. flexneri* O-antigens, *Glycobiology*, **23**, 475-485.
29. Sun, Q., Lan, R., Wang, J., Xia, S., Wang, Y., Wang, Y., Jin, D., Yu, B., Knirel, Y. A., and Xu, J. (2013) Identification and characterization of a novel *Shigella flexneri* serotype Yv in China, *PLoS ONE*, **8**, e70238.
30. Kenne, L., Lindberg, B., Petersson, K., and Romanowska, E. (1977) Basic structure of the oligosaccharide repeating-unit of the *Shigella flexneri* O-antigens, *Carbohydr. Res.*, **56**, 363-370.
31. Carlin, N. I. A., and Lindberg, A. A. (1987) Monoclonal antibodies specific for *Shigella flexneri* lipopolysaccharides: clones binding to type IV, V, and VI antigens, group 3,4 antigen, and an epitope common to all *Shigella flexneri* and *Shigella dysenteriae* type 1 strains, *Infect. Immun.*, **55**, 1412-1420.
32. Carlin, N. I. A., Bundle, D. R., and Lindberg, A. A. (1987) Characterization of five *Shigella flexneri* variant Y-specific monoclonal antibodies using defined saccharides and glycoconjugate antigens, *J. Immunol.*, **138**, 4419-4427.
33. Wang, J., Lan, R., Knirel, Y. A., Luo, X., Senchenkova, S. N., Shashkov, A. S., Xu, J., and Sun, Q. (2014) Serological identification and prevalence of a novel O-antigen epitope linked to 3- and 4-O-acetylated rhamnose III of lipopolysaccharide in *Shigella flexneri*, *J. Clin. Microbiol.*, **52**, 2033-2038.
34. Shashkov, A. S., Senchenkova, S. N., Sun, Q., Lan, R., Wang, J., Perepelov, A. V., Knirel, Y. A., and Xu, J. (2013) Structure of the O-antigen of a novel *Shigella flexneri* serotype, 1d (I: 7,8), *Carbohydr. Res.*, **373**, 93-96.
35. Kenne, L., Lindberg, B., Petersson, K., Katzenellenbogen, E., and Romanowska, E. (1977) Structural studies of the *Shigella flexneri* variant X, type 5a and 5b O-antigens, *Eur. J. Biochem.*, **76**, 327-330.
36. Dmitriev, B. A., Knirel, Y. A., Sheremet, O. K., Shashkov, A. S., Kochetkov, N. K., and Hofman, I. L. (1979) Somatic antigens of *Shigella*. The structure of the specific polysaccharide of *Shigella newcastle* (*Sh. flexneri* type 6) lipopolysaccharide, *Eur. J. Biochem.*, **98**, 309-316.
37. Wehler, T., and Carlin, N. I. A. (1988) Structural and immunochemical studies of the lipopolysaccharide from a new provisional serotype of *Shigella flexneri*, *Eur. J. Biochem.*, **176**, 471-476.
38. Foster, R. A., Carlin, N. I. A., Majcher, M., Tabor, H., Ng, L.-K., and Widmalm, G. (2011) Structural elucidation of the O-antigen of the *Shigella flexneri* provisional serotype 88-893: structural and serological similarities with *Shigella flexneri* provisional serotype Y394 (1c), *Carbohydr. Res.*, **346**, 872-876.
39. Kondakova, A. N., Vinogradov, E. V., Shekht, M. E., Markina, A. A., Lindner, B., L'vov, V. L., Aparin, P. G., and Knirel, Y. A. (2010) Structure of the oligosaccharide region (core) of the lipopolysaccharides of *Shigella flexneri* types 2a and 5b, *Bioorg. Khim.*, **36**, 396-399.
40. Kubler-Kielb, J., Vinogradov, E., Mocca, C., Pozsgay, V., Coxon, B., Robbins, J. B., and Schneerson, R. (2010) Immunochemical studies of *Shigella flexneri* 2a and 6, and *Shigella dysenteriae* type 1 O-specific polysaccharide-core fragments and their protein conjugates as vaccine candidates, *Carbohydr. Res.*, **345**, 1600-1608.
41. Guan, S., Bastin, D. A., and Verma, N. K. (1999) Functional analysis of the O antigen glucosylation gene cluster of *Shigella flexneri* bacteriophage SfX, *Microbiology*, **145**, 1263-1273.
42. Carlin, N. I. A., Wehler, T., and Lindberg, A. A. (1986) *Shigella flexneri* O-antigen epitopes: chemical and immunochemical analyses reveal that epitopes of type III and group 6 antigens are identical, *Infect. Immun.*, **53**, 110-115.
43. Lugowski, C., Romanowska, E., Kenne, L., and Lindberg, B. (1983) Identification of a trisaccharide repeating-unit in the enterobacterial common-antigen, *Carbohydr. Res.*, **118**, 173-181.
44. Dell, A., Oates, J., Lugowski, C., Romanowska, E., Kenne, L., and Lindberg, B. (1984) The enterobacterial common-antigen, a cyclic polysaccharide, *Carbohydr. Res.*, **133**, 95-104.
45. Carlin, N. I., Rahman, M., Sack, D. A., Zaman, A., Kay, B., and Lindberg, A. A. (1989) Use of monoclonal antibodies to type *Shigella flexneri* in Bangladesh, *J. Clin. Microbiol.*, **27**, 1163-1166.
46. Talukder, K. A., Dutta, D. K., Safa, A., Ansaruzzaman, M., Hassan, F., Alam, K., Islam, K. M., Carlin, N. I., Nair, G. B., and Sack, D. A. (2001) Altering trends in the dominance of *Shigella flexneri* serotypes and emergence of serologically atypical *S. flexneri* strains in Dhaka, Bangladesh, *J. Clin. Microbiol.*, **39**, 3757-3759.
47. Rusden, A. D., Stephenson, D. P., and Verma, N. K. (2013) Topological investigation of glucosyltransferase V in *Shigella flexneri* using the substituted cysteine accessibility method, *Biochemistry*, **52**, 2655-2661.
48. Markine-Goriaynoff, N., Gillet, L., Van Etten, J. L., Korres, H., Verma, N., and Vanderplasschen, A. (2004) Glycosyltransferases encoded by viruses, *J. Gen. Virol.*, **85**, 2741-2754.
49. Mavris, M., Manning, P. A., and Morona, R. (1997) Mechanism of bacteriophage SfII-mediated serotype conversion in *Shigella flexneri*, *Mol. Microbiol.*, **26**, 939-950.
50. Allison, G. E., Angeles, D., Tran-Dinh, N., and Verma, N. K. (2002) Complete genomic sequence of SfV, a serotype-converting temperate bacteriophage of *Shigella flexneri*, *J. Bacteriol.*, **184**, 1974-1987.
51. Allison, G. E., Angeles, D. C., Huan, P. T., and Verma, N. K. (2003) Morphology of temperate bacteriophage SfV and characterization of the DNA packaging and capsid genes: the structural genes evolved from two different phage families, *Virology*, **308**, 114-127.
52. Sun, Q., Lan, R., Wang, Y., Wang, J., Wang, Y., Li, P., Du, P., and Xu, J. (2013) Isolation and genomic characterization of SfI, a serotype-converting bacteriophage of *Shigella flexneri*, *BMC Microbiol.*, **13**, 39.
53. George, D. T., Stephenson, D. P., Tran, E., Morona, R., and Verma, N. K. (2013) Complete genome sequence of SfII, a serotype-converting bacteriophage of the highly prevalent *Shigella flexneri* serotype 2a, *Genome Announc.*, **1**, e00626-13.

54. Jakhetia, R., Talukder, K. A., and Verma, N. K. (2013) Isolation, characterization and comparative genomics of bacteriophage SfIV: a novel serotype converting phage from *Shigella flexneri*, *BMC Genomics*, **14**, 677.
55. Sun, Q., Lan, R., Wang, Y., Wang, J., Luo, X., Zhang, S., Li, P., Wang, Y., Ye, C., Jing, H., and Xu, J. (2011) Genesis of a novel *Shigella flexneri* serotype by sequential infection of serotype-converting bacteriophages SfX and SfI, *BMC Microbiol.*, **11**, 269.
56. Stagg, R. M., Tang, S. S., Carlin, N. I., Talukder, K. A., Cam, P. D., and Verma, N. K. (2009) A novel glucosyltransferase involved in O-antigen modification of *Shigella flexneri* serotype 1c, *J. Bacteriol.*, **191**, 6612-6617.
57. Thanweer, F., Tahiliani, V., Korres, H., and Verma, N. K. (2008) Topology and identification of critical residues of the O-acetyltransferase of serotype-converting bacteriophage, SF6, of *Shigella flexneri*, *Biochem. Biophys. Res. Commun.*, **375**, 581-585.
58. Casjens, S., Winn-Stapley, D. A., Gilcrease, E. B., Morona, R., Kuhlewein, C., Chua, J. E., Manning, P. A., Inwood, W., and Clark, A. J. (2004) The chromosome of *Shigella flexneri* bacteriophage Sf6: complete nucleotide sequence, genetic mosaicism, and DNA packaging, *J. Mol. Biol.*, **339**, 379-394.
59. Sun, Q., Lan, R., Wang, Y., Wang, J., Xia, S., Wang, Y., Zhang, J., Yu, D., Li, Z., Jing, H., and Xu, J. (2012) Identification of a divergent O-acetyltransferase gene oac_{1b} from *Shigella flexneri* serotype 1b strains, *Emerg. Microbes Infect.*, **1**, e21.
60. Sun, Q., Knirel, Y. A., Lan, R., Wang, J., Senchenkova, S. N., Shashkov, A. S., Wang, Y., Wang, Y., Luo, X., and Xu, J. (2014) Dissemination and serotype modification potential of pSFxv_2, an O-antigen PEtN modification plasmid in *Shigella flexneri*, *Glycobiology*, **24**, 305-313.
61. Zhang, N., Lan, R., Sun, Q., Wang, J., Wang, Y., Zhang, J., Yu, D., Hu, W., Hu, S., Dai, H., Du, P., Wang, H., and Xu, J. (2014) Genomic portrait of the evolution and epidemic spread of a recently emerged multidrug-resistant *Shigella flexneri* clone in China, *J. Clin. Microbiol.*, **52**, 1119-1126.
62. Ye, C., Lan, R., Xia, S., Zhang, J., Sun, Q., Zhang, S., Jing, H., Wang, L., Li, Z., Zhou, Z., Zhao, A., Cui, Z., Cao, J., Jin, D., Huang, L., Wang, Y., Luo, X., Bai, X., Wang, P., Xu, Q., and Xu, J. (2010) Emergence of a new multidrug-resistant serotype X variant in an epidemic clone of *Shigella flexneri*, *J. Clin. Microbiol.*, **48**, 419-426.
63. Qiu, S., Wang, Z., Chen, C., Liu, N., Jia, L., Liu, W., Wang, L., Hao, R., Zhang, L., Wang, Y., and Song, H. (2011) Emergence of a novel *Shigella flexneri* serotype 4s strain that evolved from a serotype X variant in China, *J. Clin. Microbiol.*, **49**, 1148-1150.