OCCASIONAL REVIEW

Structure and biology of *Shigella flexneri* O antigens

D. A. R. SIMMONS and ELZBIETA ROMANOWSKA*

*Department of Bacteriology and Immunology, University of Glasgow, Glasgow G11 6NT and *Ludwig Hirsfeld Institute of Immunology and Experimental Therapy, Polish Academy of Sciences, Wroclaw, Poland*

The species *Shigella flexneri* comprises a serologically heterogeneous group of dysentery bacilli whose O antigens are polysaccharide-lipid A-polypeptide-lipid B complexes strictly comparable in their gross structural and biological properties to the analogous antigens present in all gram-negative bacilli of the family Enterobacteriaceae (Simmons, 1971a). As shown in fig. 1, the lipopolysaccharide (LPS) component of these antigens comprises two distinct regions—a common basal structure shared by all *S. flexneri* serotypes (except serotype 6) and O-specific side-chains that determine the serological specificity and cross-reactivity of the whole antigen (Simmons, 1966). The common basal sugars 2-keto-3-deoxy-octonate, L-glycero-D-manno-heptose phosphate, D-glucose, D-galactose and N-acetyl-D-glucosamine are incorporated into the growing basal structure in that order and defects in the genes controlling this biosynthetic pathway result in an incomplete basal structure of rough (R) specificity (Johnston *et al.*, 1967). The O-specific side-chains (each consisting of six to eight repeating pentasaccharide units) may be further subdivided into two regions—a primary unbranched chain of L-rhamnose and N-acetyl-D-glucosamine identical to the variant Y antigen, and secondary side-chains of α -D-glucosyl and O-acetyl residues which are substituted on the primary chain in a position that is unique for each serotype (Simmons, 1969). The specificity of the linkage of the glucose secondary side-chain is governed by specific UDP-glucose transferases that are phage-dependent and map near the *lac* locus (Manson *et al.*, 1970). Loss of the phage attachment site in recombination experiments with *Escherichia coli* HfrC (*lac*⁺) strains usually leads to loss of the glucose secondary side-chains and, thus, to the production of lactose-fermenting hybrids of variant Y-type specificity (Manson *et al.*, 1970). However, hybrids with other specificities have also been described (Roma-

nowska and Lachowicz, 1970; Katzenellenbogen *et al.*, 1973; Lugowski *et al.*, 1975).

In an earlier review (Simmons, 1971a), structural and genetic aspects of the biosynthesis of the *S. flexneri* O antigens were considered in detail. The conclusions drawn at that time about the structural changes involved in smooth to rough (S→R) mutation, in the production of X and Y variants, and in the modification of type-specificity by lysogenic conversion remain valid today. However, some of the structures themselves require to be revised, largely as a result of methylation studies performed during the past decade by Lindberg and his colleagues (Lindberg *et al.*, 1973; Kenne *et al.*, 1977a and b; Kenne *et al.*, 1978). This review summarises the evidence for these currently accepted structures and indicates how these studies have elucidated the biology of the *S. flexneri* O antigens.

Structure of the basal region of S. flexneri lipopolysaccharides

The biosynthesis of the *S. flexneri* LPS basal region proceeds from the 2-keto-3-deoxy-octonate-lipid A core of chemotype Re by the sequential addition of L-glycero-D-manno-heptose phosphate, D-glucose, D-galactose and N-acetyl-D-glucosamine to give chemotypes Rd, Rc, Rb and Ra respectively (Johnston *et al.*, 1967). (A chemotype is a family of antigens with the same qualitative sugar composition.) The structures of all these antigens have been elucidated but most interest has centred upon the *S. flexneri* serotype 4b-R LPS because it is a chemotype Ra polymer identical with the complete basal structure. Early studies of this LPS have shown that it contains glucose, galactose and N-acetylglucosamine in the proportions 3:1:1 (Johnston *et al.*, 1967). Methylation analysis of the native polysaccharide and structural studies following N-deacetylation-deamination and Smith degradation of this polymer (Jansson *et al.*, 1979), implied that

Fig. 1. Schematic representation of the gross structure of *S. flexneri* LPS showing the constituents of the rough (R)-specific common basal structure and the smooth (O)-specific side-chains. The O-specific side-chains consist of a primary chain (of N-acetylglucosamine and rhamnose) identical with the group factor 3,4 variant Y antigen and secondary side-chains (of glucose or O-acetyl radicals or both) that determine the type specificity of the various smooth serotypes.

the basal pentasaccharide side-chain has the following structure*:

These findings were confirmed by methylation analysis and by structural analysis of a series of oligosaccharides obtained from this polymer by

partial acid hydrolysis (Simmons, 1983a). This pentasaccharide structure requires some revision of that originally proposed by Johnston *et al.* (1967) and is of special interest in that it is identical with the *E. coli* R3 basal structure (Jansson *et al.*, 1979, 1981). The structural relationship of this basal sequence to the rest of the *S. flexneri* LPS molecule is shown in fig. 2.

Structure of O-specific region of *S. flexneri* LPSs

The structures of the two regions of the O-specific side-chains—the primary unbranched variant Y chains (of L-rhamnose and N-acetyl-D-glucosamine) and the type-specific secondary side-chains (of α -D-glucosyl and O-acetyl residues)—were elucidated as follows.

Structure of the S. flexneri variant Y primary chain. Although earlier studies (Simmons, 1969) gave much information about the composition and

*Key to sugar abbreviations used in this review: Glc = glucose; Gal = galactose; Rha = rhamnose; GlcNH₂ = glucosamine; GlcNAc = N-acetylglucosamine; GalNAc = N-acetylgalactosamine; GalA = galacturonic acid; KDO = 2-keto-3-deoxy-octonate (3-deoxy-2-oxo-octonate); Hep = glyceromannoheptose; PETN = phosphoryl-ethanolamine; α and β define the anomeric form of the glycosidic linkage; D and L define the optical isomeric form. All sugar residues discussed in this review were of the pyranose configuration.

factors were shown to be related to α -D-glucosyl and O-acetyl secondary side-chains substituted on the primary (variant Y) chain as follows: factor I, α -D-glucose (1 $\rightarrow$ 4)-linked to the N-acetylglucosamine residue; factor II, α -D-glucose (1 $\rightarrow$ 4)-linked to the RhaIII residue; factor IV, α -D-glucose (1 $\rightarrow$ 6)-linked to the N-acetylglucosamine residue; factor V, α -D-glucose (1 $\rightarrow$ 3)-linked to the RhaII residue and group factor 7,8, α -D-glucose (1 $\rightarrow$ 3)-linked to the RhaI residue. Variant-X specificity is determined by the possession of group factor 7,8 only, i.e., by α -D-glucose (1 $\rightarrow$ 3)-linked to the RhaI residue of the variant Y primary chain. Specific factor III and group factor 6 are both related to the immunodominant O-acetyl group 2-O-substituted on the RhaIII residue because de-acetylation destroys both specificities (Romanowska *et al.*, 1971). The group-factor-6 activity is related to the partial structure 2-O-acetyl- α -L-Rha and the factor III specificity to the complete sequence 2-O-acetyl- α -L-Rha-(1 $\rightarrow$ 3)- β -D-GlcNAc. This conclusion is supported by the observation that substitution of the GlcNAc residue with α -D-glucose (as in serotypes 1b and 4b) masks the factor-III specificity whereas the factor-6 activity continues to be expressed. The precise determinant for group-factor 3,4 has not yet been completely elucidated but it is related to O-acetyl groups substituted on the primary (variant Y) chain. The structures assigned to the repeating units of the different *S. flexneri* O-specific side-chains are summarised in fig. 3.

Analysis of *S. flexneri* O-specific and group sero-determinants. Evidence for the molecular groupings that determine the O-specific and group factors has come from combined chemical and serological studies in which structurally-defined oligosaccharides obtained by acid hydrolysis of the *S. flexneri* O-specific polysaccharides (Simmons, 1969) were used to inhibit the homologous complement fixation systems (Simmons, 1971*b*). The factor-I system was inhibited by α -D-Glc-(1 $\rightarrow$ 4)-GlcNAc; the factor-II system by α -D-Glc-(1 $\rightarrow$ 4)-Rha; the factor-IV system by α -D-Glc-(1 $\rightarrow$ 6)-GlcNAc and the factor-V system by α -D-Glc-(1 $\rightarrow$ 3)-Rha. However, the saccharides inhibiting the factor-7,8 and -III systems are now known to be α -D-Glc-(1 $\rightarrow$ 3)-Rha and 2-O-acetyl-Rha respectively and not α -D-Glc-(1 $\rightarrow$ 2)-Rha and 2-O-acetyl-Glc as previously reported (Simmons, 1971*b*). As noted above, the complete factor-III determinant is a 2-O-acetyl- α -L-Rha-(1 $\rightarrow$ 3)- β -D-GlcNAc sequence but this specificity is lost (masked) when the GlcNAc residue is substituted with α -D-glucose. These findings not only confirm the chemical structures assigned to the different antigenic factors but provide serological evidence

that the proposed sequences are indeed the sero-determinants of the O-specific side-chains. The structures assigned to these sero-determinants are summarised in the table.

Structure of *S. flexneri* serotype-6 LPSs

S. flexneri of serotype 6 has always been regarded as atypical by many bacteriologists for several reasons but has retained its place within this genus because of its biochemical and antigenic relationships with other *S. flexneri* serotypes and, above all, because of its clinical association with acute bacillary dysentery. The more important members of this group include the Boyd 88, Manchester and Newcastle biotypes. As with all other members of the family Enterobacteriaceae, their LPSs have a basal region to which the O-specific side-chains are attached.

The basal structure of serotype 6 LPSS. Two detailed studies of the serotype-6 basal region have been performed. In the first, Katzenellenbogen and Romanowska (1980) used methylation analysis, periodate oxidation, Smith degradation and enzymic modification of core oligosaccharides to elucidate the complete basal region of a Manchester biotype (and its rough mutants) and found the following structure:

All the glycosidic linkages in this structure are of the α -type except for the side-chain glucose which is β -linked (Gamian and Romanowska, 1982). All the hexoses are of the D-configuration.

In the second of these studies (Simmons, 1983b), structural analysis of a series of oligosaccharides obtained by partial acid hydrolysis and methylation analysis of the native polysaccharide confirmed that the pentasaccharide hexose region of the basal structure of another Manchester strain was the same as that above. This structure is of particular interest because it is identical with the *E. coli* RI core which is also found in *S. sonnei* phase II (Gamian and Romanowska, 1982) and in *E. coli* C (Jansson *et al.*, 1981) strains.

Fig. 3. Comparison of the structures of various O-specific side-chain repeating units with that of a classical variant Y strain. The variant Y structure which is also elaborated by Boyd's 103B strain and all Y-type hybrids, is a cryptic situation in all smooth *S. flexneri* serotypes. Throughout this report, the antigenic formulae for the different serotypes follow the classification of Ewing and Carpenter (1966) in which type-specific factors are designated by roman capitals and group factors by arabic numerals.

Table. Molecular structures of *S. flexneri* antigenic determinants

<i>S. flexneri</i> factor	Structure of determinant*
I	α -glucosyl-(1 $\rightarrow$ 4)- β -N-acetylglucosamine
II	α -glycosyl-(1 $\rightarrow$ 4)- α -rhamnose (III)
III	2-O-acetyl- α -rhamnosyl-(1 $\rightarrow$ 3)- β -N-acetylglucosamine
IV	α -glucosyl-(1 $\rightarrow$ 6)- β -N-acetylglucosamine
V	α -glucosyl-(1 $\rightarrow$ 3)- α -rhamnose (II)
VI	β -rhamnosyl-(1 $\rightarrow$ 3)- β -N-acetylgalactosamine
3, 4	O-acetyl groups substituted on the (variant Y) primary side-chain
6	2-O-acetyl- α -rhamnose (III)
7, 8	α -glucosyl-(1 $\rightarrow$ 3)- α -rhamnose (I)

* All the sugar residues are of the D-configuration except for the rhamnose units which are of the L form. (For fuller details of structures and symbols, see text.)

The structures of the basal regions of the Boyd 88 and Newcastle biotypes have not yet been investigated.

The O-specific side-chains of serotype-6 LPSs. The only detailed study of the O-specific side-chain of a serotype-6 (biotype Manchester) LPS by modern methods is that of Katzenellenbogen *et al.* (1976) who proposed the following trisaccharide repeating unit:

Some of these repeating units have a further β -L-rhamnosyl residue substituted on the 4-O-position of the above 'non-reducing end-group' rhamnose and about 64% of the repeating units carry an O-acetyl radical substituted on one or other of the rhamnose residues. The O-acetyl and N-acetyl groups seemed to play a role in, but were not essential to, factor-VI specificity which was attributed mainly to the sequence β -L-Rha-(1 $\rightarrow$ 3)-GalNAc.

Using similar methods (including nmr spectroscopy and methylation analysis), Dmitriev *et al.* (1979) found that the repeating unit of a serotype-6 (Newcastle) O-specific polysaccharide had the following tetrasaccharide structure.

From a comparison of this structure with that of *S. flexneri* variant Y, the cross-reactivity of serotype 6

with other *S. flexneri* serotypes was assumed to be due to the common 2-O-substituted α -L-rhamnosyl residue, but this was not tested experimentally in any serological system.

Although *S. flexneri* serotype-6 strains share many features (including DNA-relatedness) with other *S. flexneri* serotypes (Brenner, 1976), the above studies show that their O-specific LPSs are structurally different. These differences are quite marked, there being no known instance in which either the basal structure or O-specific side-chain of a serotype-6 LPS has been demonstrated in any other *S. flexneri* serotype. Furthermore, if the above structures are confirmed, they would imply marked differences between the O-specific side-chains of the various serotype-6 (Manchester and Newcastle) biotypes.

Enzyme defects associated with smooth to rough mutation in S. flexneri

As indicated above, the enzyme defects responsible for smooth-to-rough (S $\rightarrow$ R) mutation are of two main types. In the first, the defect blocks the completion of the basal region so that the O-specific side-chains (which continue to be synthesised) cannot be attached to the growing molecule because of lack of the appropriate receptor. In these circumstances, the intact O-specific side-chains are found free in the ultracentrifuge supernates of the extracted rough LPSs. In contrast, the second type of defect involves a block in a synthetase or transferase for the O-specific side-chains, which are, consequently, not synthesised and, therefore, not detectable in the ultracentrifuge supernates of these rough mutants.

In an analysis of the supernatant fluids of a series of *S. flexneri* R-LPS preparations (Johnston *et al.*, 1968), those of rough serotypes 1a-R, 2a-R, 3a-R, 4a-R₁, 4a-R₂, 5-R and variant X-R contained free side-chains bound together as in the O-specific structures, because they precipitated with the homologous smooth *S. flexneri* antiserum. These findings, together with the observation that these rough mutants lacked one or more of the basal sugars, indicate a lesion in one of the basal synthetases or transferases of these mutants. By contrast, the absence of O-specific side-chains in the supernates of R-LPSs of serotypes 1b-R, 4b-R, 6-R and variant Y-R indicated the loss of a side-chain synthetase or transferase in this group of mutants. However, the 1b-R and variant Y-R mutants were exceptional in that they also had deficient basal structures and were, therefore, presumed to have a defect in a single enzyme

required for the synthesis of both regions of the molecule. However, the results did not exclude the possibility that these mutants had a double enzyme defect.

The probable nature of the enzyme defects in the different R-mutants was defined somewhat more precisely by Johnston *et al.* (1968) on the assumption that the biosynthetic pathways for the incorporation of exogenous glucose and galactose into bacterial LPSs shown in fig. 4 (Osborn *et al.*, 1964) also hold true for *S. flexneri*. They concluded as follows:

Rd mutants. These mutants elaborated a LPS identical with the heptose phosphate-KDO-lipid A 'backbone' but were unable to incorporate the first glucose residue into the basal structure side-chain. This failure may be due to their inability to synthesise uridine diphosphate (UDP)-glucose or to their inability to transfer glucose from this sugar nucleotide to heptose phosphate. With *S. flexneri* serotype 3a-R, the former possibility was excluded by the demonstration of glucose in the O-specific material in the ultracentrifuge supernate so this mutant probably lacks the UDP-glucose transferase analogous to the UDP-glucose transferase I of salmonellae (Osborn *et al.*, 1964). With *S. flexneri* serotype 1b-R, the absence of both O-specific side-chain and basal glucose indicates that this mutant is unable to synthesise UDP-glucose. Fig. 4 shows that such a failure could result from a defect in glucokinase, phosphoglucumutase or UDP-glucose pyrophosphorylase but, as this mutant synthesised non-specific polyglucosan in which the first two enzymes are necessary for transglucosylation (Sigal *et al.*, 1964), the defective enzyme is believed to be UDP-glucose pyrophosphorylase.

Rc mutants. The Rc LPSs of serotypes 1a-R, 2a-R, 4a-R₁, 5-R and variant X-R resembled those of salmonellae in that all contained heptose phosphate and glucose but these mutants were unable to incorporate galactose into the growing side-chain of the basal structure when grown in the presence of exogenous glucose. This defect in galactose incorporation was by-passed in mass-cultures in the presence of exogenous galactose because the mu-

tants then synthesised complete smooth O-specific LPS. Because Rc mutants contain glucose in their LPS, they must possess all the enzymes necessary to synthesise UDP-glucose (fig. 4). Their failure to produce smooth LPS could be due to a defect in UDP-galactose-4-epimerase or UDP-galactose transferase. However, the production of smooth LPS from exogenous galactose rules out the possibility of a defect in the latter enzyme, so these Rc mutants are probably deficient in UDP-galactose-4-epimerase.

Rb mutants. Two mutants of this type (serotype 4b-R₂ and variant Y-R) contained heptose phosphate, glucose and galactose in their polysaccharides and N-acetyl-glucosamine in their lipid-A component. Because they were, therefore, able to synthesise all the basal sugars, they were presumed to be unable to incorporate the N-acetylglucosamine residue that completes the basal structure for lack of the appropriate UDP-N-acetyl-glucosamine transferase.

Ra mutants. Mutants of this type complete the basal structure but lack O-specific side-chains which contain glucose, N-acetylglucosamine and rhamnose. Because these mutants were able to synthesise both glucose and N-acetylglucosamine for their basal structure, they are presumed to lack a rhamnose synthetase or an O-specific side-chain transferase.

Semi-rough (SR) mutant. From quantitative structural and serological analysis, the LPS of *S. flexneri* serotype 4a-R₃ mutant was shown to be a semi-rough polymer consisting of the basal structure and the first repeating unit of the O-specific side-chain. Such polymers are an intermediate stage in the biosynthesis of the smooth LPS. Mutants of this type, which were first described by Naide *et al.* (1965), were assumed to possess transferase I which transfers the first repeating unit of the O-specific side-chain to the basal structure but not the transferase II system which transfers the second repeating unit to the first. The *S. flexneri* SR mutant is presumed to be defective in the analogous transferase II.

Fig. 4. Biosynthetic pathways for the incorporation of exogenous glucose and galactose into bacterial LPS (after Osborn *et al.*, 1964). 1. glucokinase; 2. phosphoglucumutase; 3. UDP-glucose pyrophosphorylase; 4. UDP-glucose transferase; 5. UDP-galactose-4-epimerase; 6. galactokinase; 7. galactose-1-P uridyl transferase; 8. UDP-galactose transferase.

The finding that the above R-antigens have overlapping structures and that identical R-LPSs can be isolated from R-mutants isolated from different smooth serotypes (Johnston *et al.*, 1967), is strong evidence that the proposed basal structure is common to all smooth *S. flexneri* serotypes (with the notable exception of *S. flexneri* serotype 6).

Comparison of the basal LPS structures of S. flexneri and other Enterobacteriaceae

There is much evidence to support the view that the basal structures of the LPS of most Enterobacteriaceae consist of the five basal sugars, 2-keto-3-deoxy-octonate, heptose phosphate, glucose, galactose and N-acetylglucosamine. However, some *E. coli* serotypes, including O 17, O 44, O 59 and O 77 have no galactose (Kauffmann *et al.*, 1960; Jann, 1965); the *E. coli* basal structures of types R1 and R4 have no N-acetylglucosamine (Jansson *et al.*, 1981), and *S. dysenteriae* (shiga) has little or no glucose (Davies *et al.*, 1955; Davies, 1957; Simmons, 1957). Thus the basal regions of most Enterobacteriaceae LPSs are qualitatively similar but clearly not identical. Structural studies confirm this view and indicate that the vast majority of the basal region side-chains that have been character-

ised to date belong to one of the six types summarised in fig. 5 (Jann and Jann, 1984). From this figure it will be apparent that the common *S. flexneri* basal region is identical with that of *E. coli* R3 (Simmons, 1983a) and that the exceptional *S. flexneri* serotype 6 structure is identical with that of *E. coli* R1 (Simmons, 1983b). Of the other *Shigella* types, the LPS of *S. sonnei* phase II also belongs to *E. coli* type R1 (Gamian and Romanowska, 1982) but that of *S. dysenteriae* (shiga) appears to differ from all previously characterised basal regions in being glucose deficient.

A further point of comparative interest lies in the internal structures of salmonellae and *S. flexneri*, which are identical in the regions corresponding to the Rc antigen. The possession of this common structure was confirmed by the observation that the Rc LPSs of *S. flexneri* and some *Salmonella* spp. (*S. typhimurium* tmM and *S. minnesota* mR5 which are also known to have an α -D-Gal-(1 $\rightarrow$ 3)- α -D-Glc determinant) are serologically indistinguishable throughout the regions of antibody excess, equivalence and antigen excess in quantitative precipitin-type-curves obtained in homologous and heterologous complement fixation systems (Simmons, 1971a). These results imply that the internal basal regions of salmonella and *S. flexneri* LPSs must be

Fig. 5. Comparison of the basal structure of LPS of salmonellae with other allied genera of the family Enterobacteriaceae (after Jann and Jann, 1984). The common *S. flexneri* basal structure is identical with that of *E. coli* R3 and the exceptional *S. flexneri* serotype-6 basal structure with that of *E. coli* R1. (The N-acetylglucosamine residue in *E. coli* K12, the terminal 'non-reducing' glucose in *E. coli* R1 and the terminal galactosyl 1 $\rightarrow$ 4 residue in *E. coli* R4 are in β -glycosidic linkage. All other residues are α -glycosidically-linked and all the hexoses are of the D-form.)

to an antigen-carrier-lipid (ACL). Polymerisation of these activated sugar repeating units is then effected by a specific polymerase in a 'zip-fastener-like' process under the direction of the *rfe* gene. This *rfe*-independent pattern of biosynthesis, which can be inhibited by bacitracin, is characterised by the formation of solvent-extractable, multiple repeating units during in-vitro synthesis and by the formation of semi-rough (SR) mutants such as *S. flexneri* serotype 4a-R₃. In contrast, the *rfe*-dependent biosynthetic system is not inhibited by bacitracin, does not produce solvent-extractable, multiple repeating units during in-vitro synthesis and does not give rise to SR mutants. The *S. flexneri* O-specific primary side-chains that are synthesised by this *rfe*-independent system are structurally and serologically identical with the variant Y antigen.

The third stage of biosynthesis is that of postpolymerisation modification of the primary side-chain. With *S. flexneri*, as with some salmonellae (Jann and Jann, 1984), this involves transglucosylation and trans-O-acetylation of the primary chain. This modification which is induced by regulating temperate bacteriophages, is described in more detail below.

The fourth and final stage of antigen synthesis involves the attachment of the completed O-specific side-chain to the basal structure. This process of translocation is carried out by the enzyme translocase under the direction of the two genes *rfaL* and *rfbT* which specifically recognise the basal structure and O-specific side-chain respectively (Jann and Jann, 1984). A schematic representation of the biosynthesis of the *S. flexneri* smooth and rough

The diagram illustrates the biosynthetic pathway of the R-series, showing the progression from Re to Ra and then to variant Y, which branches into 1a, 2a, 4a, and variant X (which further branches into 3a and 5b). Below the main pathway, a detailed breakdown of the sugar components and the enzymes involved in their synthesis is provided. The sugar components are: Heptose, Glucose, Galactose, N-acetylglucosamine, Rhamnose or N-acetylglucosamine, and Glucose. The enzymes involved are: Heptose synthetase or transferase, UDP-glucose pyrophosphorylase or UDP-glucose transferase, UDP-galactose 4-epimerase, N-acetylglucosamine transferase, Rhamnose or N-acetylglucosamine transferase or rhamnose synthetase, and UDP-glucose transferase. The diagram also shows the genetic control of the pathway, with genes rfa, rfb, rfaL, and rfbI controlling the biosynthetic genes, and specific phage genes controlling the UDP-glucose transferase.

Chemotype/serotype (R-series rough; others smooth)

Re → Rd → Rc → Rb → Ra → variant Y → 1a, 2a, 4a, variant X → 3a, 5b

Sugar components and Enzyme defects:

- Heptose: Sugar not incorporated; Enzyme defect: Heptose synthetase or transferase
- Glucose: Enzyme defect: UDP-glucose pyrophosphorylase or UDP-glucose transferase
- Galactose: Enzyme defect: UDP-galactose 4-epimerase
- N-acetylglucosamine: Enzyme defect: N-acetylglucosamine transferase
- Rhamnose or N-acetylglucosamine: Enzyme defect: Rhamnose or N-acetylglucosamine transferase or rhamnose synthetase
- Glucose: Enzyme defect: UDP-glucose transferase

Genetic Control:

- Controlling genes: rfa, rfb, rfaL, rfbI
- Specific phage genes: 3a transacetylase or UDP-glucose transferase

Fig. 6. Structural relationships, biosynthetic pathways and the genetic control of *S. flexneri* smooth and rough antigens showing the probable nature of the sugars and enzymes involved in the biosynthetic blocks that give rise to the various mutants and variants.

antigens is given in fig. 6. It shows the probable nature of the sugars and enzymes that are involved in the biosynthetic blocks that give rise to the different serotypes and mutants.

Taxonomic aspects of S. flexneri O-antigens

Past difficulties with the serology and taxonomy of *S. flexneri* have been attributed to poor reproducibility of results and to considerable uncertainty as to the interpretation of the data. In particular, the taxonomic relationship of certain serotypes and the status of X and Y variants have been the subject of much debate that has led to the introduction of many different classifications based on one or other of two concepts.

In the first of these (Andrewes and Inman, 1919), it was suggested that the serological behaviour of *S. flexneri* was explained by postulating four main races designated V, W, X and Z. Each serotype was a mixture of these four factors with one in quantitative predominance. The Y race was a more even mixture of the four factors, but the inability of suspensions of the four races to absorb Y antiserum led the authors to conclude that they had "failed to solve the problem of the antigenic structure of the Y races". The second concept was introduced by Boyd (1938) who proposed six main serotypes, each with a type-specific antigen that never occurred in other serotypes and a group antigen that was shared. X and Y strains were regarded as variants, Y being a degraded form of many types and X a degraded form of Z (serotype 3a). Boyd also believed that A→B variation represented the loss of type specific antigen with the production of a pure group variant which had been isolated on three occasions—strains 103B, 119B and the original variant Y of Hiss and Russell (1903). This concept of antigenic structure has been the basis of all subsequent classifications including those of the Shigella Commission Reports (1951, 1958) and that of Ewing and Carpenter (1966).

In our view, the structural studies reviewed above are of interest and relevance to these concepts in the following respects. First, they show that the antigenic structure of *S. flexneri* serotype is neither a mixture of type-specific antigens nor a mixture of one type-specific with group antigens, but is a single molecular entity carrying different antigenic determinants. This view is confirmed by the isolation of small glucose-containing oligosaccharides of undoubted purity that carry both type and group specificity. Second, Boyd's view that each serotype contains one type-specific and one or more group factors is confirmed by our structural results in that

type-specificity is determined by the unique linkage of the terminal α -glucosyl secondary side-chain and group specificity by the presence of an identical structural sequence in the primary chain of all serotypes. Third, the suggestion that group antigen was present in most serotypes has been substantiated by our observation that the structures of Boyd's original 103B strain and Y-specific *S. flexneri* × *E. coli* hybrid LPSs from all serotypes (Manson *et al.*, 1970) are identical with that of classical variant Y strains. However, the LPS from the original variant Y of Hiss and Russell (1903) which many observers thought closer to W (serotype 2a), has proved to be structurally and serologically identical with serotype-2a LPS (Simmons, 1971a). Fourth, the view that variant X is a degraded form is consistent with the observation that it is a de-acetylated serotype 3a. Fifth, the structural analysis of the *S. flexneri* LPSs is relevant to the debate that has surrounded the serological status and biological significance of X and Y variants. The variant Y LPS (an intermediate stage in the normal biosynthesis of the complete smooth antigen) is elaborated as the end-product when there is a loss or repression of the gene controlling the UDP-glucose transferase necessary to incorporate the type specific, α -glucosyl secondary side-chains. The fact that it is incomplete, agrees with the taxonomist's view that it is a degraded type-specific antigen. The nature of the X variant is still somewhat equivocal. Its determinant belongs to the same α -glucosyl family as those of the type-specific serotypes and the gene directing its UDP-glucose transferase maps genetically with the type specific or T-locus near the *lac* region. On the other hand, it may also be regarded as a degraded form in the sense that it is an intermediate stage in the biosynthesis of serotypes 3a and 5b from the precursor Y structure. These findings offer an explanation for the difficulties that have arisen in the past in interpreting the status of X variants. Sixth, the observation that *S. flexneri* O antigens all consist of a variant Y structure substituted with an immunodominant α -glucosyl or O-acetylated residue, explains to some extent the poor reproducibility of results obtained by early workers. With such closely related, cross-reacting polymers, the antigenic differences between the various *S. flexneri* serotypes are much less distinct than in most other genera and their serotyping that much more difficult.

One of the objectives of these studies was the elucidation of the molecular basis of serological specificity and cross-reactivity to improve our understanding of taxonomic problems. This aim has been achieved in so far as the general structure

of the *S. flexneri* O-antigens has been established, the precise structural sequence of specific and group antigens elucidated, the relationship of serotype 3a and variant X determined and the nature and significance of X and Y antigens clarified.

Genetic aspects of S. flexneri O-antigens

In an earlier review, Simmons (1971a) proposed that the *S. flexneri* O antigens were a closely related family of polymers with primary side-chains of pattern Y₁ or Y₂, to which the α -glucosyl secondary side-chains that determine serological specificity were attached. However, further structural studies (Lindberg *et al.*, 1973) and the isolation of one type of variant Y from all serotypes, prove rather conclusively that there is but one Y-type structure which is common to all *S. flexneri*. As these findings seem to be well-established, it should be possible to recover Y-type variants from all *S. flexneri* strains when the genes controlling the UDP-glucose transferases are lost or defective. Naturally occurring examples of this kind of mutation may be found in the early literature (Hiss, 1904), but detailed studies of the phenomenon have been described on three occasions only. In the first of these, Kruse *et al.* (1907) found that the Wilz B strain derived from Wilz A (serotype 1a) was serologically indistinguishable from Strong B (a classical variant Y). In the second detailed study of A→B variation, Boyd (1931) noted that *S. flexneri* type 103A frequently mutated to a stable form, 103B, which he believed to be closely related to, but not quite identical with, variant Y. Our studies have shown it to be structurally identical with variant Y. In the third study of A→B mutation, Ewing (1954) isolated a classical Y variant from a serotype 2a strain. All these observations show that Y-type variants arise in nature as a consequence of A→B variation and that the B form and variant Y are structurally and serologically identical.

More recently, Y-type variants have been produced experimentally in the laboratory by conjugating *S. flexneri* serotypes with *E. coli* Hfr strains. The first experiments of this type were performed by Luria and Burrous (1957) who isolated recombinants of this class from serotypes 2a, 2b and 4a and showed that the genes controlling type specificity mapped near the *lac* locus. Even more recently, it has been shown (Timakov *et al.*, 1970) that Y-type hybrids can be obtained from all *S. flexneri* serotypes by use of *E. coli* HfrC as the donor strain and that the type specific or T-loci for antigens I, II, IV, V and 7,8 (but not III, VI or 6) also map near the *lac* locus. These findings not only imply a common Y-

type structure in all *S. flexneri* strains but show that the T-locus controls precisely those determinants in which glucose is the immunodominant sugar whereas in antigens III and 6, which map elsewhere, the immunodominant structures are related to O-acetyl-rhamnose residues. These structural and genetic findings, taken together with the observation that the loss of type specificity is accompanied by the synthesis of a glucose-deficient Y-type structure, imply that the T-locus is a gene or cluster of genes controlling the synthesis of the specific UDP-glucose transferases required to complete the biosynthesis of the various *S. flexneri* LPSs from the precursor Y-structure.

Even more recently, changes in *S. flexneri* O-specificity have been shown to depend on phage or lysogenic conversion. Matsui (1958) converted antigens I and III of serotypes 1a and 3a, respectively, to antigen IV with a phage from serotype 4c. Using the appropriate phages, other workers converted the analogous antigens of 1a, 1b, 2a, 3b and variant Y to IV (Iseki and Hamano, 1959) and those of 2a, 2b, 4a, variant X and variant Y to I (Okada *et al.*, 1960). A variant Y strain was doubly converted to variant X and then serotype 2b by Giammanco (1968) with the two appropriate phages f-7,8 and f-II sequentially. From the molecular structures shown in fig. 3, it will be apparent that these phage conversions produce very simple differences in the linkage of the glucose secondary side-chains. Indeed, they could all be accomplished by a single enzyme change. Until recently, the precise role of these type-specific phages had been poorly understood, there being some doubt as to whether the phages themselves carry in the regulating genes or simply act as de-repressors allowing the expression of pre-existing genes in the bacterial chromosome. However, the studies of Petrovskaya *et al.* (1972) have shown that stable hybrids of *E. coli* K12 carrying the *S. flexneri* variant Y character *his*⁺ 3,4⁺ gene, were converted to serotypes 5 and variant X by the appropriate V and 7,8 phages. This result provides good evidence that the regulating genes are carried in by the phage because it is most improbable that the *E. coli* K12 genome codes for specific *S. flexneri* antigens. Moreover, there is also good genetic evidence (Petrovskaya and Nevskaya, 1975) that the 7,8 prophage, which also attaches near the *lac* locus, does so by a different method from the other type-specific phages. These findings not only provide evidence for the role of prophages in determining type specificity but also offer an explanation for the results obtained in the hybridisation experiments described above. In recombination with *E. coli*

K12, the replacement of the *lac* region deletes the adjacent phage attachment site so that the phage genes controlling the type specific UDP-glucose transferases cannot be incorporated into the bacterial genome. The biosynthesis of the O antigen is, therefore, halted at the group-specific variant Y stage. The locations of the various genes controlling O-antigen biosynthesis have been provisionally mapped on the *S. flexneri* chromosome by Petrovskaya and Licheva (1982) as shown in fig. 7.

The results of these lysogenic and conjugation studies seem to be of particular interest and

importance in that they have elucidated some aspects of the mechanisms involved in the genetic control of *S. flexneri* O-antigen synthesis which may now be summarised as follows. The biosynthesis of these molecules commences with the KDO-lipid A "backbone" of chemotype Re and proceeds to the complete basal structure of chemotype Ra under the control of *rfa* genes as described above. The repeating units of the variant Y primary side-chain are assembled separately on the antigen-carrier lipid (ACL) under the direction of *rfb* genes and polymerised under the direction of the *rfc* gene.

Fig. 7. Provisional linkage map of *S. flexneri* chromosome (after Petrovskaya and Licheva, 1982) showing the position of loci involved in the biosynthesis of *S. flexneri* smooth and rough antigens.

Post-polymerisation modification by specific O-glucosylation or O-acetylation of the primary chain then proceeds under the control of the specific phage genes. The synthesis of the whole antigen is finally completed by attachment of the O-specific side-chains to the basal structure by translocase under the control of the genes *rfaL* and *rfbT*.

Conclusion

In this review, an attempt has been made to correlate the structure and biology of *S. flexneri* O-antigens. These studies are of intrinsic value in that they have solved the molecular basis of serological specificity and cross-reactivity, elucidated the biochemical and enzymic basis of smooth to rough

(S→R) mutation and A→B variation, defined the structural relationship and biosynthetic pathways of smooth and rough antigens, and indicated the mechanisms involved in the genetic control of *S. flexneri* O-antigen biosynthesis. It seems reasonable to predict that similar studies with other organisms will continue to reveal structural and genetic relationships between different genera that will ultimately clarify the evolutionary history of Enterobacteriaceae.

We acknowledge with thanks our indebtedness to Professor B. Lindberg of the Arrhenius Laboratory, Department of Organic Chemistry, University of Stockholm, Sweden for helpful advice and constructive criticism in the preparation of this review.

REFERENCES

- Andrewes F W, Inman A C 1919 A study of the serological races of the Flexner group of dysentery bacilli. Medical Research Committee, Special report series, London, No. 42.
- Boyd J S K 1931 Some investigations into so-called "non-agglutinable" dysentery bacilli. *Journal of the Royal Army Medical Corps* **57**: 161–186.
- Boyd J S K 1938 The antigenic structure of the mannitol-fermenting group of dysentery bacilli. *Journal of Hygiene* **38**: 477–499.
- Brenner D J 1976 *Shigella flexneri* 6 biotypes: a review. *Health Laboratory Science* **13**: 218–222.
- Bundle D R, Josephson S 1980 Artificial carbohydrate antigens: the synthesis of a tetrasaccharide hapten, A *Shigella flexneri* O-antigen repeating unit. *Carbohydrate Research* **80**: 75–85.
- Davies D A L 1957 Isolation of a "rough" somatic antigen from *Shigella dysenteriae*. *Biochimica et Biophysica Acta* **26**: 151–158.
- Davies D A L, Morgan W T J, Record B R 1955 Studies in immunochemistry. 15. The specific polysaccharide of the dominant 'O'-somatic antigen of *Shigella dysenteriae* (Shiga). *Biochemical Journal* **60**: 290–303.
- de Belder A N, Norrman B 1968 The distribution of substituents in partially acetylated dextran. *Carbohydrate Research* **8**: 1–6.
- Dmitriev B A, Knirel Y A, Sheremet O K, Shashkov A A, Kochetkov N K, Hofman I L 1979 Somatic antigens of *Shigella*. The structure of the specific polysaccharide of *Shigella newcastle* (*Sh. flexneri* type 6) lipopoly-saccharide. *European Journal of Biochemistry* **98**: 309–316.
- Ewing W H 1954 Mannitol negative varieties of *Shigella flexneri* serotypes. *Journal of Immunology* **72**: 404–410.
- Ewing W H, Carpenter K P 1966 Recommended designations for the subserotypes of *Shigella flexneri*. *International Journal of Systematic Bacteriology* **16**: 145–149.
- Gamian A, Romanowska E (1982) The core structure of *Shigella sonnei* lipopolysaccharide and the linkage between O-specific polysaccharide and the core region. *European Journal of Biochemistry* **129**: 105–109.
- Giammanco G 1968 Conversion lysogénique de caractères antigéniques de *Shigella flexneri* (antigène II de type et complexe 7,8 de groupe). *Annales de l'Institut Pasteur* **114**: 63–76.
- Hiss P H 1904 On fermentation and agglutinative characters of bacilli of the "dysentery group". *Journal of Medical Research* **13**: 1–51.
- Hiss P H, Russell F F 1903 A study of a bacillus resembling the bacillus of Shiga from a case of fatal diarrhea in a child, with remarks on the recognition of dysentery, typhoid and allied bacilli. *Medical News, New York* **82**: 289–295.
- Iseki S, Hamano S 1959 Conversion of type antigen IV in *Shigella flexneri* by bacteriophage. *Proceedings of the Japanese Academy* **35**: 407–412.
- Jann B 1965 Vergleichende untersuchungen an spezifischen polysacchariden von *Escherichia coli* (O und K antigene). Thesis. Universität Freiburg, Germany.
- Jann K, Jann B 1984 Structure and biosynthesis of O-antigens. In: Rietschel E T (ed) In: Handbook of endotoxin, vol. 1. Elsevier, Amsterdam, pp 138–186.
- Jansson P E, Lindberg A A, Lindberg B, Wollin R 1981 Structural studies on the hexose region of the core in lipopolysaccharides from Enterobacteriaceae. *European Journal of Biochemistry* **115**: 571–577.
- Jansson P E, Lindberg B, Lindberg A A, Wollin R 1979 Structural studies on the hexose region of the Enterobacteriaceae type R3 core polysaccharide. *Carbohydrate Research* **68**: 385–389.
- Johnston J H, Johnston R J, Simmons D A R 1967 The immunochemistry of *Shigella flexneri* O-antigens. The biochemical basis of smooth to rough mutation. *Biochemical Journal* **105**: 79–87.
- Johnston J H, Johnston R J, Simmons D A R 1968 The immunochemistry of *Shigella flexneri* O-antigens: a study of smooth to rough mutation. *Archivum Immunologiae et Therapiae Experimentalis* **16**: 252–259.
- Josephson S, Bundle D R 1980 Artificial carbohydrate antigens: synthesis of rhamnose trisaccharide and disaccharide haptens common to *Shigella flexneri* O-antigens. *Journal of the Chemical Society Perkin I* 297–301.
- Katzenellenbogen E, Lugowski C, Mulczyk M, Romanowska E 1973 Immunochemical studies on *Shigella flexneri* 5b serotype and its recombinant. *Archivum Immunologiae et Therapiae Experimentalis* **21**: 699–702.
- Katzenellenbogen E, Mulczyk M, Romanowska E 1976 Structural studies on O-specific polysaccharide of *Shigella flexneri* serotype 6. *European Journal of Biochemistry* **61**: 191–197.
- Katzenellenbogen E, Romanowska E 1980 Structural studies on

- Shigella flexneri* serotype 6 core region. *European Journal of Biochemistry* **113**: 205–211.
- Kauffmann F, Braun O H, Lüderitz O, Stierlin H, Westphal O 1960 Zur immunchemie der O-antigene von Enterobacteriaceae. IV. Analyse der zuckerbausteine von Escherichia O-antigenen. *Zentralblatt für Bacteriologie, 1 Abteilung Originale* **180**: 180–188.
- Kenne L, Lindberg B, Petersson K, Katzenellenbogen E, Romanowska E 1977a Structural studies of the *Shigella flexneri* variant X, type 5a and type 5b O-antigens. *European Journal of Biochemistry* **76**: 327–330.
- Kenne L, Lindberg B, Petersson K, Romanowska E 1977b Basic structure of the oligosaccharide repeating unit of the *Shigella flexneri* O-antigens. *Carbohydrate Research* **56**: 363–370.
- Kenne L, Lindberg B, Petersson K, Katzenellenbogen E, Romanowska E 1978 Structural studies of *Shigella flexneri* O-antigens. *European Journal of Biochemistry* **91**: 279–284.
- Kruse, Ritterhaus, Kemp, Metz 1907 Dysenterie und pseudodysenterie. *Zeitschrift für Hygiene und Infektionskrankheiten* **52**: 417–438.
- Lindberg B, Lönngren J, Ruden U, Simmons D A R 1973 A reinvestigation of *Shigella flexneri* O-antigens by methylation analysis. *European Journal of Biochemistry* **32**: 15–18.
- Lüderitz O, Staub A M, Westphal O 1966 Immunochemistry of O and R antigens of *Salmonella* and related Enterobacteriaceae. *Bacteriological Reviews* **30**: 192–255.
- Lugowski C, Katzenellenbogen E, Romanowska E, Mulczyk M 1975 Immunochemical studies on the serotype 2b of *Shigella flexneri* and its recombinant. *Archivum Immunologiae et Therapiae Experimentalis* **23**: 305–308.
- Luria S E, Burrous J W 1957 Hybridization between *Escherichia coli* and *Shigella*. *Journal of Bacteriology* **74**: 461–476.
- Manson R, Simmons D A R, Petrovskaya V G 1970 Antigenic changes in *Shigella flexneri* following conjugation with *Escherichia coli* HfrC. *European Journal of Biochemistry* **17**: 472–476.
- Matsui S 1958 Antigenic changes in the *Shigella flexneri* group by bacteriophage. *Japanese Journal of Microbiology* **2**: 153–158.
- Naide Y, Nikaido H, Mäkelä P H, Wilkinson R G, Stocker B A D 1965 Semi-rough strains of *Salmonella*. *Proceedings of the National Academy of Sciences of the USA* **53**: 147–153.
- Okada M, Hoshin T, Karate T 1960 Lysogenic conversion of somatic antigens of *Shigella flexneri* by bacteriophage. *Science (Tokyo)* **30**: 90 (in Japanese. English resumé in *Excerpta Medica* section 4, 1962, vol. 15, p 664).
- Osborn M J, Rosen S M, Rothfield L, Zeleznick L D, Horecker B L 1964 Lipopolysaccharide of the gram-negative cell wall. Biosynthesis of a complex heteropolysaccharide occurs by successive addition of specific sugar residues. *Science* **145**: 783–789.
- Petrovskaya V G, Licheva T A 1982 A provisional chromosome map of *Shigella* and the regions related to pathogenicity. *Acta Microbiologica Academiae Scientiarum Hungaricae* **29**: 41–53.
- Petrovskaya V G, Nevskaya N A 1975 Effect of some prophages on genetic behaviour of *Shigella flexneri*. *Genetika* **11**(10): 61–66 (in Russian).
- Petrovskaya V G, Nevskaya N A, Khomenko N A 1972 Reproduction of phage conversion of the somatic antigen of *Shigella flexneri* on hybrids of *Escherichia coli*. *Genetika* **8**(9): 113–118 (in Russian).
- Romanowska E, Lachowicz T M 1970 Studies on O-specific polysaccharides of *Shigella flexneri* 1b (I) and its smooth mutant Z (III). *FEBS letters* **8**: 293–296.
- Romanowska E, Katzenellenbogen E, Lugowski C, Mulczyk M 1971 The chemical nature of *Shigella flexneri* group antigen factor 6. *Pathologie et Microbiologie (Basel)* **37**: 352–360.
- Shigella Commission Report 1951 Nomenclature and classification of the dysentery bacilli. *International Bulletin of Bacteriology, Nomenclature and Taxonomy* **1**: 150–166.
- Shigella Commission Report 1958 Supplement to the third report on the Shigella group. *International Bulletin of Bacteriology, Nomenclature and Taxonomy* **8**: 93–95.
- Sigal N, Cattaneo J, Segal I H 1964 Glycogen accumulation by wild-type and uridine diphosphate glucose pyrophosphorylase-negative strains of *Escherichia coli*. *Archives of Biochemistry and Biophysics* **108**: 440–451.
- Simmons D A R 1957 The hexose constituents of some *Shigella* polysaccharide hydrolysate. *Journal of General Microbiology* **17**: 650–657.
- Simmons D A R 1966 The immunochemistry of *Shigella flexneri* lipopolysaccharides. A quantitative analysis of their monosaccharide constituents. *Biochemical Journal* **98**: 903–908.
- Simmons D A R 1969 The immunochemistry of *Shigella flexneri* O-antigens. The structure and biosynthesis of the O-specific side chains of some representative serotypes. *European Journal of Biochemistry* **11**: 554–575.
- Simmons D A R 1971a Immunochemistry of *Shigella flexneri* O-antigens: a study of structural and genetic aspects of the biosynthesis of cell-surface antigens. *Bacteriological Reviews* **35**: 117–148.
- Simmons D A R 1971b The immunochemistry of *Shigella flexneri* O-antigens: an analysis of the determinants of some representative type specific factors. *Immunology* **20**: 17–24.
- Simmons D A R 1981 Structure of the O-specific lipopolysaccharides from *Shigella flexneri* variant Y and from Y-type hybrids of *Shigella flexneri* × *Escherichia coli* HfrC. *Biochemical Society Transactions* **9**: 554–555.
- Simmons D A R 1983a Structure of the basal region of *Shigella flexneri* lipopolysaccharides. *Biochemical Society Transactions* **11**: 104–105.
- Simmons D A R 1983b Structure of *Shigella flexneri* serotype 6-R (rough) lipopolysaccharide. *Biochemical Society Transactions* **11**: 386–387.
- Timakov V D, Petrovskaya V G, Bondarenko V M 1970 Studies on the genetic control of *Shigella* subgroup B type specific antigens. 1. Behaviour of *Shigella flexneri* type-specific antigens in sexual recombination of *Shigella* × *E. coli*. *Annales de l'Institut Pasteur* **118**: 3–9.