Published in final edited form as:
Virology. 2007 December 20; 369(2): 234–244. doi:10.1016/j.virol.2007.07.027.

The Genome of ε15, a Serotype-Converting, Group E1 *Salmonella enterica*-Specific Bacteriophage

Andrew M. Kropinski[a,b], **Irina V. Kovalyova**[a,c], **Stephen J. Billington**[d], **Aaron N. Patrick**[e], **Brent D. Butts**[e], **Jared A. Guichard**[e], **Trevor J. Pitcher**[e], **Carly C. Guthrie**[e], **Anya D. Sydlaske**[e], **Lisa M. Barnhill**[e], **Kyle A. Havens**[e], **Kenneth R. Day**[e], **Darrel R. Falk**[e], and **Michael R. McConnell**[e,*]

a *Department of Microbiology and Immunology, Queens University, Kingston, Ontario K7L 3N6, Canada*

b *Public Health Agency of Canada, Laboratory for Foodborne Zoonoses, Guelph, Ontario N1G 3W4, Canada*

c *Biology Department, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada*

d *Department of Veterinary Science and Microbiology, University of Arizona*

e *Department of Biology, Point Loma Nazarene University, San Diego, CA 92106*

Abstract

The genome sequence of the *Salmonella enterica* serovar Anatum-specific, serotype-converting bacteriophage ε15 has been completed. The nonredundant genome contains 39,671 bp and 51 putative genes. It most closely resembles the genome of ϕV10, an *Escherichia coli* O1H57-specific temperate phage, with which it shares 36 related genes. More distant relatives include the *Burkholderia cepacia*-specific phage, BcepC6B [8 similar genes], the *Bordetella bronchiseptica*-specific phage, BPP-1 [8 similar genes] and the *Photobacterium profundum* prophage, P Pϕpr1 [6 similar genes].

ε15 gene identifications based on homologies with known gene families include the terminase small and large subunits, integrase, endolysin, two holins, two DNA methylase enzymes (one adenine-specific and one cytosine-specific) and a RecT-like enzyme. Genes identified experimentally include those coding for the serotype conversion proteins, the tail fiber, the major capsid protein and the major repressor. ε15's attP site and the Salmonella attB site with which it interacts during lysogenization have also been determined.

Keywords

bacteriophage; serotype conversion; virion proteins; Epsilon 15 genome; attP; attB

Introduction

Bacteriophage ε15 is a Group E1 *Salmonella enterica* serovar Anatum-specific, serotype-converting phage that belongs to the order *Caudovirales* ("tailed viruses") and the family *Podoviridae* (phages with short, non-contractile tails; Ackermann, 1999). During the 1950s,

*Submitting author. Mailing address: Dr. Michael R. McConnell, Department of Biology, Point Loma Nazarene University, San Diego, CA 92106, Tel:619-849-2304, Fax: 619-849-2598, Email: E-mail: mmcconne@pointloma.edu.

ε15 was shown to orchestrate a serological change in its host cell, a phenomenon that Salvadore Luria, Hisao Uetake and their co-investigators called “cell surface conversion” (Uetake, et al., 1958; Uetake, et al., 1955; Uetake & Uchida, 1959). Although a novelty at the time, cell surface conversion was eventually shown to be commonplace among temperate bacteriophages (Uetake, 1979).

During the 1960s, Phil Robbins and his collaborators at MIT defined the chemical structure of the Group E1 *Salmonella enterica* O-polysaccharide and showed that the serological changes brought about by ε15 during cell surface conversion involved replacement of this O-polysaccharide polymer, comprised of D-Mannosyl-β1→4-L-Rhamnosyl-α1→3-D-O-Acetyl-Galactose repeat units joined together by α1→6 glycosidic linkages, with a non-acetylated polymer of the same repeat unit, held together by β1→6 glycosidic bonds (Bray & Robbins, 1967; Losick and Robbins, 1967; Robbins, et al, 1965; Robbins & Uchida, 1962; Robbins & Uchida, 1965; Uchida, et al, 1963). They further concluded that ε15 engineered the change in O-polysaccharide structure by producing: 1) an inhibitor protein that blocks the activity of the host cell O-polysaccharide alpha polymerase enzyme; 2) an O-polysaccharide beta polymerase enzyme that replaces the inhibited host cell alpha polymerase; and 3) a repressor protein that blocks transcription of the host cell O-polysaccharide acetyltransferase gene, thereby causing acetylation of galactose residues to be gradually curtailed (Losick & Robbins, 1969).

Bacteriophage ε15 drew major attention again in the early 1970’s when researchers in Boston and Kyoto independently discovered that its tail spikes possess endorhamnosidase activity capable of degrading Group E1 *S. enterica* O-polysaccharide polymers down to D-O-Acetyl-Galactosyl-α1→6-D-Mannosyl-β1→4-L-Rhamnose end-products (Kanegasaki & Wright, 1973; Takeda & Uetake, 1973). Subsequent biochemical and genetic studies with ε15 mutants indicated that the purpose of the tail spike endorhamnosidase activity is to bring the ε15 virion into closer proximity with the surface of the outer membrane prior to release of its DNA (McConnell, et al., 1979). Several other phages specific for smooth (i.e. O-polysaccharide-containing) *Enterobacteriaciae* were subsequently shown to have tail fibers that enzymatically degrade the O-polysaccharide portions of their respective host cell LPS molecules (see review by Wright, et al., 1980).

We report here a detailed analysis of the ε15 genome, including its probable transcriptional regulatory regions and its DNA sequences that function during the lysogenization process. Several of its genes have been identified by experimentation, including those coding for the tail fiber, the major capsid protein, the repressor and four cell surface conversion proteins. Other workers have recently reported on ε15 virion proteins that comprise the DNA packaging/injection apparatus (Jiang, et al, 2006) and their work is discussed in light of our own. The genome of ε15 displays little similarity to the genomes of other known *Salmonella* phages, but instead, is most closely related to that of ϕV10, an *Escherichia coli* O1H57-specific temperate phage, with which it shares 36 related genes.

Results and Discussion

General Features of the Genome

Earlier restriction analysis of ε15 DNA had shown that the genome is circularly permuted, with a mass of approximately 40.3 kb and a terminal redundancy of ~0.6 kb (McConnell, et al., 1992). The DNA sequence data supports that finding, in that the non-redundant sequence contains 39,671 basepairs. To circumvent the problems in graphically representing a circularly-permuted, terminally-redundant genome, the ε15 genomic map (Fig. 1) was arbitrarily opened adjacent to a 17-bp sequence (**CCGCCG**ACTAT**GGCGGC**TTTGTTTT), located just upstream of the probable gene for the small terminase subunit (designated as gene *1* on the

map). All other genes were numbered sequentially in a clockwise manner from this point and are described in Table 1.

The overall base composition of ε15 DNA (50.85 mol% GC) resembles that of *Salmonella* species (52 mol% GC), as does the pattern of its codon usage, with only a few minor exceptions involving codons specifying proline, threonine and lysine. A scan of the entire genome, using a window of 100 bp revealed several regions of higher AT content, including the regions corresponding to the cell surface conversion genes (*21, 22* and *28)*, the attP site between genes 26 and 27, and the region between genes *38* and *39* (Figure 1*)*.

A Grigoriev GC/AT-skew analysis of the ε15 genome revealed an AT-skew maximum and a corresponding GC minimum within *gene 41* (34,150 +/−100bp) suggesting that this may be where DNA replication originates (Grigoriev, 1998; Grigoriev, 1999; Kowalczuk, et al., 2001; Lobry, 1999). *Gene 41* encodes a predicted cytosine-specific methyltransferase (one of two methyltransferase genes, the other being gene *29*, which encodes an adenine-specific enzyme). While the product of gene *42* exhibits sequence similarity to bacterial primosomal protein, no other gene products resembling known DNA replication proteins are present, suggesting that ε15 is reliant on one or more host cell enzymes for replication of its genome.

Our analysis indicates that gene *38* encodes the major repressor protein. Near-consensus, diverging promoters and operator-like sequences positioned on each side of gene 38 probably regulate expression of ε15's immediate early genes (Figure 1), with delayed early and late expression most likely dependent upon terminator read-through mechanisms. Although the protein product of gene *38* exhibits poor sequence similarity with other known phage repressors, it is similar in size (198 amino acids) and it contains a helix-turn-helix motif (pfam01381; smart00530) of the type that typically serves as the operator recognition element for repressor proteins (amino acids M104 through G150, 21 of which are good H-bonders). Finally, the prototype clear plaque mutant of ε15 (known as ε15*vir* in the literature) contains a single altered base pair in gene *38* (an A/T > G/C change at base pair 203), which results in a D68 to G68 change in the mutant protein (McConnell, unpublished data). As with the repressor genes of Lambda, D3 (Kropinski, 2000) and phage r1t (van Sinderen, et al., 1996), ε15 gene *38* lacks an identifiable RBS.

An analysis of the genome suggests that the tail spike gene and lysis genes are probably the last to be expressed during an infection. Late expression of the lysis genes could explain ε15's relatively long latent period (~55 minutes) and large burst size (~300 PFUs/cell) at 37° C (McConnell, unpublished data). There appear to be three lysis proteins, altogether (Table 1). *Genes 23 and 24* both code for holins and *gene 25* codes for an endolysin, based upon a variety of shared characteristics with known lysis proteins produced by other bacteriophages (Grundling, et al., 2000;Liu, et al., 2004;Ramanculov and Young, 2001;Young, 1992;Young and Blasi, 1995). The property of having two linked genes that both code for holins has also been observed for *Streptococcus thermophilus* phage, φO1205 (Sheehan, et al., 1999), and *Lactococcus lactis* phage BK5-T (Mahanivong, et al., 2001).

Integration

Southern Blot studies performed at PLNU during the 90's involving restriction endonuclease-digested S. anatum (ε15) lysogen DNA suggested that the attP site of ε15 was located near the phage's integrase gene (Orf27). This has now been confirmed by "chromosome walking", using *S.* anatum (ε15) DNA and fimers (Fidelity Systems, Inc.) bracketing a non-coding region of the ε15 genome positioned between genes 26 and 27. The results, depicted in Figure 2, reveal two regions of homology ("common core" regions) between the ε15 and host cell genomes: 1) a 13 base pair segment encompassing the last four codons of the Salmonella guaA gene wherein recombination occurs (the proximity of prophage ε15 to the gua locus of Salmonella

was first suggested long ago by the conjugation studies of Matsuyama and Uetake (1972)); and 2) a nearby 31 base pair, AT-rich segment which appears to be non-coding within both the ε15 and host cell genomes. Surrounding the common core sequences and spanning 558 base pairs of the ε15 genome altogether, are four copies of a direct repeat sequence (GTGACGGTAT) that probably represent "arm-type" binding sequences for the integrase. Also present are two likely IHF binding sites, one of which overlaps the segment in which recombination occurs (see legend to Figure 2).

The ε15 integrase is clearly a tyrosine recombinase, as indicated both by the structural features of its attP site and by the presence of all six appropriately-positioned amino acid residues that are highly conserved in this family of proteins, including the catalytic tyrosine at position Y_{370} (see review by Groth and Calos, 2004).

Morphogenesis

Our earlier work indicated that ε15 packages DNA by a head-full packaging mechanism (i.e. use of a terminase complex that initiates packaging at a *pac* site; (McConnell, et al., 1992)). Sequence homology analyses now reveal that genes 1 and 2 code for ε15's small and large terminase subunits, respectively. Immediately downstream of the terminase genes, where one might expect to see genes specifying the portal (or head-tail connector) protein, we find instead a gene (*2A)* that is transcribed in the opposite direction from the remainder of the putative morphogenesis genes. Gene *2A* encodes a protein which is homologous to gene 66 of *Shigella* phage, Sf6 (Casjens, et al., 2004).

Figure 3 and Table 2 depict the results of protein composition studies on ε15 virions (510S) and virion-like, DNA-less particles that move with an S value of ~173 ± 12S on 5–40% sucrose sedimentation gradients. All of the pI values presented in Table 2 were obtained by 2D analyses of the 173S particles, since their proteins are more amenable to isoelectric focusing, due to the absence of DNA.

Several virion structural proteins have been matched experimentally with their corresponding genes in the ε15 genome, one being V-1, which comprises the enzymatically active tail parts of ε15. In the 70's, it was shown that non-infectious, non-adsorbing, virion-like particles formed by an ε15 nonsense mutant called *am*2 were normal-looking under the electron microscope, except for the absence of tail parts; furthermore, when these particles were analyzed on SDS/polyacrylamide gels, only the V-1 polypeptide was missing (McConnell, et al., 1979). The experimentally measured size of 115,508 ± 2,626 amu for V-1 shown in Table 2 is very close to the inferred size of gp*20*, which is 115,616 amu (Table 1). To further confirm this identification, we determined the gene *20* sequence of am2 and three other tail part-deficient, missense mutants of ε15 known to map in the same gene as am2, based upon earlier *in vivo* complementation analyses (McConnell, 1976). All four mutants contain a single, unique mutational change in gene 20 and there is colinearity between the physical positions of the mutational alterations and their genetic map positions, as determined by earlier two- and three-factor genetic recombination experiments (data not shown).

Investigators at MIT and Baylor College of Medicine have recently used mass spectrometry measurements on tryptic digest fragments of ε15 virion proteins extracted from SDS-PA gels to independently confirm that the tail fiber protein is gp20 and that the protein products of genes *4*, 7, *11*, *15* and *17* are also present in mature ε15 virions, with gp4 and gp7 likely being the portal protein and the major capsid protein, respectively (Jiang, et al., 2006). Our molecular weight and pI measurements on proteins contained within ε15 virions and 173S particles both confirm and extend their findings (see Table 2).

The virion-like, 173S particles contain some proteins that are not found in virions. Based upon staining with Coomassie Brilliant Blue, the most abundant of these is PV-1, which is present at ~129 copies per particle, assuming 415 capsid proteins (Figure 3, Table 2). PV-1's experimentally measured mass and pI values (~55,054 daltons and ~5.1) are similar only to those inferred for gp2 (ε15's terminase large subunit), which are 55,517 daltons and 5.4, respectively (Table 1). Much less abundant at ~17 copies per 173S particle, is PV-3, whose experimentally measured size and pI values (26,978 ± 192 daltons and 4.61 ± 0.24, respectively) closely resemble those of the gp8 scaffolding protein of bacteriophage P22 (Eppler, et al, 1991) as well as those of the inferred gene product of ε15 gene 6 (25,740 daltons and pI 4.6; Table 1). Further efforts are underway to determine why probable maturation proteins gp2 and gp6 co-purify in such abundance with 173S particles on buoyant density and sedimentation velocity gradients.

Serotype Conversion

Our three-pronged, microbiological, immunological and biochemical screening approach for the detection of the serotype converting genes of ε15 indicates that: 1) gene *21* codes for the O-polysaccharide beta polymerase enzyme; 2) gene *22* codes for the protein that inhibits the host cell O-polysaccharide alpha polymerase enzyme; and 3) genes *28* and *46* both code for proteins that can prevent acetylation of galactose residues in the Group E1 O-polysaccharide (Table 3).

Genetic confirmation that gene *21* codes for the beta polymerase was achieved by PCR amplifying and sequencing this gene, both from ε15 mutant NC5, the original beta polymerase nonsense mutant isolated by Lynn Silver in the 70's (Silver, 1975), as well as from six beta polymerase mutants newly isolated in our laboratory by screening among ε15-lysogenized *Salmonella* bacteria for mutants that had become spontaneously resistant to bacteriophage ε34, a virus that recognizes beta-linked O-polysaccharide as its receptor. All seven ε15 mutants with defective beta polymerase activity displayed a single base pair change in gene *21* (data not shown).

A comparison of the O-polysaccharide beta polymerase enzymes coded for by ε15 gene *21* and the cryptic beta polymerase gene (Orf17.4) that resides at the downstream end of the rfb gene cluster in Group E1 Salmonellae (McConnell, et al, 2001), reveals two enzymes that are similar in terms of their sizes, pI values and numbers of membrane-spanning helices (390 amino acids, pI = 9.26 and ten membrane spanning helices for ε15 gp21, versus 367 amino acids, pI = 9.5 and nine membrane spanning helices for the S. enterica *Orf17.4* gene product). Despite their physical similarities and their identical catalytic activities (both convert lipid-linked D-Mannosyl-β1→4-L-Rhamnosyl-α1→3-D-O-Acetyl-Galactose trisaccharide repeat units into β1→6 glycosidically-linked polymers), these two enzymes display little or no similarity at the primary sequence level.

We were initially assuming that the other cell surface conversion proteins of ε15 would be soluble. Losick had reported that the O-polysaccharide alpha polymerase inhibitor was a heat-resistant, water-soluble protein (Losick, 1969) and Robbins and his coworkers had shown with *in vitro* assays involving sonicates of ε15-infected cells that O-polysaccharide acetyltransferase activity levels off very soon after infection, an outcome they assumed was the result of inhibition of transcription of the acetyltransferase gene by a phage encoded repressor (Robbins, et al, 1965). We now know that the alpha polymerase inhibitor is gp22, a small protein with 66 amino acids, two predicted membrane-spanning helices and a pI value of 8.2. Protein gp*22* physically resembles the alpha-polymerase inhibitor (iap) of *Pseudomonas aeruginosa* phage D3 (Newton, et al., 2001), which has been demonstrated experimentally to be a membrane protein.

ε15 proteins gp28 and gp46 are both able to block acetylation of galactose residues in the Group E1 O-polysaccharide and TMHMM analyses indicate that both are membrane-associated (Figure 4). Gp28 was confirmed experimentally to be a membrane protein by tagging its N-terminus with Hexa-His, then using SDS-PAGE and Western Blotting to show that it co-purified with the membrane fraction following disruption of Salmonellae bacteria by sonication (data not shown). These results argue strongly against the transcriptional repression model for inhibition of galactose acetylation by ε15. Our current model is that proteins gp28 and gp46 act instead as inhibitors, but only of newly-synthesized acetyltransferase enzymes, perhaps by preventing them from orienting properly within the membrane, relative to the other enzymes that are involved in synthesis of the O-polysaccharide repeat unit.

Table 3 presents only indirect evidence that ε15 proteins gp28 and gp46 block acetylation of O-polysaccharide; namely, that when either of these two genes is placed into *S. enterica* serovar Anatum bacteria, the transformed cells remain ε15-sensitive but become resistant to g341 (g341 only infects Group E1 Salmonellae strains whose O-polysaccharide contains acetylated galactose residues). We verified that gp28 affects lipopolysaccharide (LPS) structure by: 1) purifying LPS from transformed *S. enterica* A1 cells carrying ε15 gene 28 and from the non-transformed parent strain; 2) normalizing the concentrations of the two LPS concentrations on the basis of their rhamnose contents; and 3) comparing the abilities of the two LPS preparations to inactivate g341 and ε15 phage during incubation in 10mM Tris-1mM magnesium sulfate buffer (pH 7) at 37°C. Although both preparations inactivated ε15 effectively, only the LPS from the parent strain lacking gene 28 was able to inactivate phage g341 (McConnell, unpublished data).

The closest known relative of ε15 is coliphage ϕV10 (GenBank accession number NC_007804), which is reported to carry an acetyltransferase gene, though no experimental data has yet been presented on its ability to seroconvert. We compared the conversion modules of ε15 and ϕV10 by reducing the corresponding GenBank gbk files to only include ε15 genes *19–24* and ϕV10 genes *24–28*, then aligned these segments using Mauve (Darling, et al., 2004). Homologous genes *19*(ε15) and *24*(ϕV10), as well as *23*,*24*(ε15) and *27*,*28*(ϕV10), served as anchors in the alignment (Figure 5). The alignments indicated that the tail fiber genes of the two phages differ significantly in length and that their homology resides only within the N-terminal coding portion. This region of the tail fiber is probably associated with base plate attachment, rather than receptor interaction, since a similar region has also been observed with the T7-like phages (Kovalyova & Kropinski, 2003). There is no evidence for a ϕV10 gene related to ε15 gene *22* (the alpha polymerase inhibitor), nor is there any homology between the beta-polymerase of ε15 (gp21) and the proposed acetyltransferase of ϕV10, though structurally both of them contain 10 transmembrane domains (Kall, et al., 2004; Sonnhammer, et al., 1998; Kahsay, et al., 2005). All three programs used to detect membrane spanning helices indicate that the N-terminus of the ε15 beta-polymerase is periplasmic, whereas the N-terminus of the putative ϕV10 acetyltransferase is cytoplasmic. The ϕV10 product contains a COG3274 domain (uncharacterized protein conserved in bacteria) and its closest relatives are hypothetical proteins from *Azoarcus* sp. EbN1 [YP_157885] and *Bacillus anthracis* [YP_022389]. The ε15 beta-polymerase (gp21) lacks conserved domains but is distantly related to hypothetical proteins of *Clostridium perfringens* [YP_695054] and *C. thermocellum* [ZP_00503873]. It is only on iterated BlastP analysis (Altschul, et al., 1997) that gp21 shows similarity to other O-polysaccharide polymerases.

Taxonomic position of ε15

Phage ε15's proteins display only limited homology to those of other known *Salmonella* phages, including those of g341, another temperate, Group E1 Salmonella-specific, serotype-converting Podoviridae phage whose genome has recently been sequenced at Point Loma

(McConnell, unpublished data). Instead, of the 51 potential ε15 gene products that were identified, 36 displayed significant sequence similarity (18% to 94% identity) with proteins of the *Escherichia coli* O1H57-specific, temperate phage ϕV10 (Table 1). Other podoviral genomes showing homology to that of ε15 are: *Burkholderia cepacia* phage BcepC6B [8 genes, (Summer, et al., 2006)], *Bordetella bronchiseptica*-specific phage, BPP-1 [8 genes (Liu, et al., 2004)], and *Photobacterium profundum* prophage PϕPpr1 [6 genes, (Vezzi, et al., 2004)]. For the latter three phages, it is primarily the morphogenic genes which display homology to genes of ε15.

Although NCBI currently describes ε15 as an "unclassified P22-like virus", only two of its proteins exhibit strong sequence similarity to P22 proteins (gp44 and gp45 resemble the Eae and Ead proteins of P22, respectively; Table 1). We recommend that bacteriophages ε15 and ϕV10 be considered by the International Committee on Taxonomy of Viruses for separate classification as a new genus (van Regenmortel, et al., 2000).

Materials & Methods

Phage and Bacterial Strains

All phage and bacterial strains utilized in this study came originally from the laboratory of Dr. Andrew Wright (Tufts University, Boston, MA)

Cloning and sequencing procedure—Approximately 95% of the ε15 genome was cloned into pUC18/19Cam plasmids as a collection of overlapping and/or abutting restriction endonuclease fragments (McConnell, et al., 1992). The pUC18/19Cam plasmids were originally provided by Masaki Hayashi of UCSD and are identical to pUC18/19, except that the beta-lactamase gene has been replaced by the chloramphenicol transacetylase gene. Regions of the ε15 genome that resisted cloning were ultimately bridged and sequenced using PCR. Most of the sequence was obtained using dye-tagged dideoxyribonucleotides and automated sequencers at the University of Arizona Genomic Analysis and Technology Core Facility and the San Diego State University Microchemical Core Facility.

Identification of the ε15 *attP* and *S.* anatum *attB* Sites—Lysogen DNA was provided to Fidelity Systems Inc. (Gaithersburg, MD 20879–4117, USA http://www.fidelitysystems.com), where fimers were designed for "walking" downstream from gene *26* and upstream from gene 27. The resulting, mostly bacterial DNA sequence data was compared with the gene 26/27 region of the E15 genome using LALIGN and the homologous regions that were discovered were further analyzed by comparison against the non-redundant nucleotide database at NCBI, using BLASTn.

Sequence Analysis and Definition of Genes—The DNA sequence was scanned through a 100 bp window for base compositional variation using "DNA base composition analysis tool" (http://molbiol-tools.ca/Jie_Zheng/). Potential integration host factor (IHF)-binding sites were assessed using MacTargsearch (Goodrich, et al., 1990) while potential transcriptional terminators were assessed using the GCG program "Terminator," and the Microsoft Windows software program GeSTer (Unniraman, et al., 2002). Promoter sequences were detected using Softberry's BPROM program at http://www.softberry.com/berry.phtml?topic=promoter.

Most genes (~80%) were identified either 1) experimentally; 2) by their homology with other known phage genes; or 3) by using GeneMark.hmm for Prokaryotes at http://opal.biology.gatech.edu/GeneMark/gmhmm2_prok.cgi (Lukashin & Borodovsky, 1998). Criteria used to define other genes included: (a) the presence of 30 or more codons, (b) an upstream sequence displaying similarity to the consensus ribosome-binding site [RBS,

TAAGGAGGT, (Shine & Dalgarno, 1974; Shine & Dalgarno, 1975)], and (c) either ATG or GTG as the initiation codon.

A compendium of online tools (http://molbiol-tools.ca) was employed in the analysis of the putative proteins, including: BLASTP (Altschul, et al., 1990; Altschul & Koonin, 1998), ALIGN (http://xylian.igh.cnrs.fr/bin/align-guess.cgi), and TMHMM (Sonnhammer, et al., 1998).

Nucleotide sequence accession number—The GenBank accession number for the genome of phage ε15 is AY150271 (NC_004775).

Proteomics—ε15 virion and virion-like particles lacking DNA were purified from confluent lysis agar plates using differential centrifugation, followed by a combination of CsCl buoyant density and 5–40% sucrose sedimentation velocity ultracentrifugation steps. S-values were estimated using P22 virions (510S), ϕX174 virions (113–114S) and beta-galactosidase (19S) as comparators. Protein and DNA contents of purified particles were measured using Lowry and diphenylamine colorimetric assays, respectively (Lowry, et al, 1951; Burton, 1956). The polypeptides of purified particles were resolved by electrophoresis on 16% and 10–20% Tricine/SDS/polyacrylamide gels (Invitrogen), then made visible by staining of the gels with Coomassie Brilliant Blue and quantified using the Kodak Digital Science 1D imaging system and SigmaScan Pro software. Polypeptide stoichiometries were estimated assuming 415 capsid proteins per particle, a number recently confirmed experimentally for ε15 virions by Jiang, et al (2006).

Isoelectric focusing was performed by boiling concentrated preparations of purified DNA-less, virion-like 173S particles in water for ten minutes, then mixing them one part to four parts with sample buffer containing 9.8M urea, 4% Tween 20 and 50mM dithiothreitol in order to re-solubilize their heat-denatured polypeptides. Samples were applied to BIO-RAD ReadyStrip IPG Strips (either pH 3–10 or pH 4–7) and electrofocused, using the BIO-RAD PROTEAN IEF Cell. Strips containing focused proteins were treated with SDS and iodoacetamide, then subjected to electrophoresis in the second dimension in 16% Tricine/SDS/polyacylamide gels that were afterwards stained with either silver or Coomassie Brilliant Blue.

Identification of genes involved in serotype conversion—Suspected conversion genes of ε15 were amplified using PCR primer pairs containing engineered restriction endonuclease cut sites that allowed for their properly oriented insertion into the mcs regions of pUC18/19Cam plasmids. The plasmids were electroporated into *Salmonella enterica* strains already transformed by pREP4 (Groger, et al., 1989; Invitrogen, Carlsbad, CA) and displaying LPS phenotypes appropriate for the detection of conversion gene function (plasmid pREP4 specifies kanamycin-resistance and also contains the Lac I repressor gene, thus affording greater control over the expression of cloned gene products within the transformants). Doubly transformed Salmonellae strains were induced with 1mM IPTG and characterized for expression of conversion genes using three previously described methods; namely: 1) determining their sensitivities to phages ε15, ε34, g341 and Felix O1; 2) whole-cell ELISA assays using commercially available (Difco) anti-O10 (alpha-linkages) and anti-O15 (beta-linkages) antisera; and 3) analysis of their LPS molecules following resolution by SDS-PAGE electrophoresis, oxidation by periodate and visualization by silver-staining (McConnell, Oakes, et al, 2001).

Acknowledgments

IK wishes to thank the Department of Microbiology and Immunology, the School of Graduate Studies and Research of Queen's University, and AK for financial support. AK acknowledges research funding from the Natural Science

and Engineering Research Council of Canada and extends special thanks to Professor V. Nagaraja and Ms. Ranjana Prakash (Department of Microbiology & Cell Biology, Indian Institute of Science, Bangalore) for their help with GeSTer.

MM wishes to thank Peter Reeves, who helped him produce the first ε15 DNA sequence data during a 1994 sabbatical in Sydney; Glenn Songer, who provided strong, ongoing support of the sequencing effort thereafter; Ben Fane and Hans-Wolfgang Ackermann, for information regarding the S-value of PhiX174 virions; and Andrew Wright, in whose laboratory he first became acquainted with the interesting world of *Salmonella* bacteriophages. Finally, MM thanks the Howard Hughes Research Institute, the NIH (AREA Grant 1R15GM52696-01), the NSF (RUI Grant# DMB-8608480), Research Associates of PLNU (an alumni support group) and the PLNU Administration for financing his summer undergraduate research program over the past two decades.

Reference List

Ackermann HW. Tailed bacteriophages: the order *Caudovirales*. Advances in Virus Research 1999;51:135–201. [PubMed: 9891587]

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of Molecular Biology 1990;215:403–410. [PubMed: 2231712]

Altschul SF, Koonin EV. Iterated profile searches with PSI-BLAST--a tool for discovery in protein databases. Trends in Biochemical Sciences 1998;23:444–447. [PubMed: 9852764]

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 1997;25:3389–4022. [PubMed: 9254694]

Bray D, Robbins P. Mechanism of E15 conversion studies with bacteriophage mutants. Journal of Molecular Biology 1967;30:457–475. [PubMed: 4970574]

Burton K. A study of the conditions and mechanisms of the diphenylamine reaction for the colorimetric estimation of deoxyribonucleic acid. Biochemistry 1956;62:315–323.

Casjens S, Winn-Stapley DA, Gilcrease EB, Morona R, Kuhlewein C, Chua JE, Manning PA, Clark AJ. The Chromosome of *Shigella flexneri* Bacteriophage Sf6: Complete Nucleotide Sequence, Genetic Mosaicism, and DNA Packaging. Journal of Molecular Biology 2004;339:379–394. [PubMed: 15136040]

Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Research 2004;14:1394–1403. [PubMed: 15231754]

Eppler K, Wyckoff E, Goates J, Parr R, Casjens S. Nucleotide sequence of the bacteriophage P22 gene required for DNA packaging. Virology 1991;183:519–538. [PubMed: 1853558]

Goodrich JA, Schwartz ML, McClure WR. Searching for and predicting the activity of sites for DNA binding proteins: compilation and analysis of the binding sites for *Escherichia coli* integration host factor (IHF). Nucleic Acids Research 1990;18:4993–5000. [PubMed: 2205834]

Grigoriev A. Analyzing genomes with cumulative skew diagrams. Nucleic Acids Research 1998;26:2286–2290. [PubMed: 9580676]

Grigoriev A. Strand-specific compositional asymmetries in double-stranded DNA viruses. Virus Research 1999;60:1–19. [PubMed: 10225270]

Groger RK, Morrow MD, Tykocinski ML. Directional antisense and sense cDNA cloning using Epstein-Barr virus episomal expression vectors. Gene 1989;81:285–294. [PubMed: 2478421]

Groth AC, Calos MP. Phage Integrases: Biology and Applications. Journal of Molecular Biology 2004;335:667–678. [PubMed: 14687564]

Grundling A, Bläsi U, Young R. Biochemical and genetic evidence for three transmembrane domains in the class I holin, lambda S. Journal of Biological Chemistry 2000;275:769–776. [PubMed: 10625606]

Jiang W, Chang J, Jakana J, Weigele P, King J, Chiu W. Structure of epsilon15 bacteriophage reveals genome organization and DNA packaging/injection apparatus. Nature 2006;439:612–616. [PubMed: 16452981]

Kahsay R, Liao L, Gao G. An improved hidden Markov model for transmembrane protein topology prediction and its application to complete genomes. Bioinformatics 2005;21:1853–1858. [PubMed: 15691854]

Kall L, Krogh A, Sonnhammer EL. A combined transmembrane topology and signal peptide prediction method. Journal of Molecular Biology 2004;338:1027–1036. [PubMed: 15111065]

Kanegasaki S, Wright A. Studies on the mechanism of phage adsorption: Interaction between Epsilon 15 and its cellular receptor. Virology 1973;52:160–173. [PubMed: 4803392]

Kovalyova IV, Kropinski AM. The complete genomic sequence of lytic bacteriophage gh-1 infecting *Pseudomonas putida*-evidence for close relationship to the T7 group. Virology 2003;311:305–315. [PubMed: 12842620]

Kowalczuk M, Mackiewicz P, Mackiewicz D, Nowicka A, Dudkiewicz M, Dudek MR, Cebrat S. DNA asymmetry and the replicational mutational pressure. Journal of Applied Genetics 2001;42:553–577. [PubMed: 14564030]

Liu M, Gingery M, Doulatov SR, Liu Y, Hodes A, Baker S, Davis P, Simmonds M, Churcher C, Mungall K, Quail MA, Preston A, Harvill ET, Maskell DJ, Eiserling FA, Parkhill J, Miller JF. Genomic and genetic analysis of *Bordetella* bacteriophages encoding reverse transcriptase-mediated tropism-switching cassettes. Journal of Bacteriology 2004;186:1503–1517. [PubMed: 14973019]

Lobry JR. Genomic landscapes. Microbiology Today 1999;26:164–165.

Losick R. Isolation of a trypsin-sensitive inhibitor of O-antigen synthesis involved in lysogenic conversion by bacteriophage epsilon-15. Journal of Molecular Biology 1969;42:237–246. [PubMed: 5803297]

Losick R, Robbins PW. Mechanism of ε15 conversion studied with a bacterial mutant. Journal of Molecular Biology 1967;30:445–455. [PubMed: 4970573]

Losick R, Robbins PW. The receptor site for a bacterial virus. Scientific American 1969;221:120–124. [PubMed: 5344268]

Lowry O, Rosebrough N, Farr A, Randall R. Protein measurement with the Folin phenol reagent. Journal of Biological Chemistry 1951;193:265–275. [PubMed: 14907713]

Lukashin A, Borodovsky M. GeneMark.hmm: a new solution for gene finding. Nucleic Acids Res 1998;26:1107–1115. [PubMed: 9461475]

Mahanivong C, Boyce JD, Davidson BE, Hillier AJ. Sequence analysis and molecular characterization of the *Lactococcus lactis* temperate bacteriophage BK5-T. Applied & Environmental Microbiology 2001;67:3564–3576. [PubMed: 11472933]

Matsuyama T, Uetake H. Chromosomal locations of Salmonella conversion phages: Mapping of prophages g341, ε15 and ε34 in Salmonella anatum. Virology 1972;49:359–367. [PubMed: 4559685]

McConnell, MR. PhD Dissertation. Tufts University; Boston, MA: 1976. Multiple steps are involved in the irreversible attachment of bacteriophage ε15 to its host cell.

McConnell M, Walker B, Middleton P, Chase J, Owens J, Hyatt D, Gutierrez H, Williams M, Hambright D, Barry M Jr. Restriction endonuclease and genetic mapping studies indicate that the vegetative genome of the temperate, *Salmonella*-specific bacteriophage, epsilon 15, is circularly-permuted. Archives of Virology 1992;123:215–221. [PubMed: 1312823]

McConnell MR, Oakes KA, Patrick AN, Mills DM. Two functional O-polysaccharide polymerase *wzy (rfc)* genes are present in the rfb gene cluster of Group E1 *Salmonella enterica* serovar Anatum. FEMS Microbiology Letters 2001;199:235–240. [PubMed: 11377873]

McConnell MR, Reznick A, Wright A. Studies on the initial interactions of bacteriophage Epsilon 15 with its host cell, *Salmonella anatum*. Virology 1979;94:10–23. [PubMed: 35879]

Newton GJ, Daniels C, Burrows LL, Kropinski AM, Clarke AJ, Lam JS. Three-component-mediated serotype conversion in *Pseudomonas aeruginosa* by bacteriophage D3. Molecular Microbiology 2001;39:1237–1247. [PubMed: 11251840]

Ramanculov E, Young R. An ancient player unmasked: T4 rI encodes a t-specific antiholin. Molecular Microbiology 2001;41:575–583. [PubMed: 11532126]

Robbins PW, Keller JM, Wright A, Bernstein RL. Enzymatic and kinetic studies on the mechanism of O-antigen conversion by bacteriophage E15. Journal of Biological Chemistry 1965;240:384–390. [PubMed: 14253440]

Robbins PW, Uchida T. Studies on the chemical basis of the phage conversion of O-antigens in the E-group *Salmonellae*. Biochemistry 1962;1:323–335. [PubMed: 14492698]

Robbins PW, Uchida T. Chemical and macromolecular structure of O-antigens from *Salmonella anatum* strains carrying mutants of bacteriophage E15. Journal of Biological Chemistry 1965;240:375–383.

Sheehan MM, Stanley E, Fitzgerald GF, van Sinderen D. Identification and characterization of a lysis module present in a large proportion of bacteriophages infecting *Streptococcus thermophilus*. Applied & Environmental Microbiology 1999;65:569–577. [PubMed: 9925584]

Shine J, Dalgarno L. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proceedings of the National Academy of Sciences of the United States of America 1974;71:1342–1346. [PubMed: 4598299]

Shine J, Dalgarno L. Terminal-sequence analysis of bacterial ribosomal RNA. Correlation between the 3′-terminal-polypyrimidine sequence of 16-S RNA and translational specificity of the ribosome. European Journal of Biochemistry 1975;57:221–230. [PubMed: 809282]

Silver, L. PhD Dissertation. Tufts University; Boston MA.: 1975. Studies on ε15 beta polymerase, a bacteriophage coded membrane protein.

Sonnhammer, ELL.; von Heijne, G.; Krogh, A. A hidden Markov model for predicting transmembrane helices in protein sequences. In: Glasgow, J.; Littlejohn, T.; Major, F.; Lathrop, R.; Sankoff, D.; Sensen, C., editors. Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology. AAAI Press; Menlo Park, CA: 1998. p. 175-182.

Summer EJ, Gonzalez CF, Bomer M, Carlile T, Morrison W, Embry A, Kucherka AM, Lee J, Mebane L, Morrison WC, Mark L, King MD, LiPuma MJ, Vidaver AK, Young R. Divergence and mosaicism among virulent soil phages of the *Burkholderia cepacia* complex. Journal of Bacteriology 2006;188:255–268. [PubMed: 16352842]

Takeda K, Uetake H. *In vitro* interaction between phage and receptor lipopolysaccharide: A novel glycosidase associated with phage Epsilon 15. Virology 1973;52:148–159.

Uchida T, Robbins PW, Luria SE. Analysis of the serologic determinant groups of the Salmonella E-Group O-antigens. Biochemistry 1963;2:663–668. [PubMed: 14075095]

Uetake H, Luria SE, Burrous JW. Conversion of somatic antigens in *Salmonella* by phage infection leading to lysis or lysogeny. Virology 1958;5:68–91. [PubMed: 13519750]

Uetake H, Nakagawa T, Akiba T. The relationship of bacteriophage to antigenic changes in group E Salmonellas. Journal of Bacteriology 1955;69:571–579. [PubMed: 14381379]

Uetake H, Uchida T. Mutants of *Salmonella* ε15 with abnormal conversion properties. Virology 1959;9:495–505. [PubMed: 13840183]

Uetake, H. The origin of conversion genes. In: Chakravarty, M., editor. Molecular Basis of Host/Virus Interactions. Science Press; Princeton, USA: 1979. p. 365-377.

Unniraman S, Prakash R, Nagaraja V. Conserved economics of transcription termination in eubacteria. Nucleic Acids Research 2002;30:675–684. [PubMed: 11809879]

van Regenmortel, MHV.; Fauquet, CM.; Bishop, DHL.; Carstens, EB.; Estes, MK.; Lemon, SM.; Maniloff, J.; McGeoch, DJ.; Pringle, CR.; Wickner, RB. Seventh Report of the International Committee on the Taxonomy of Viruses. Academic Press; New York: 2000. Virus Taxonomy: Classification and Nomenclature of Viruses .

van Sinderen D, Karsens H, Kok J, Terpstra P, Ruiters MH, Venema G, Nauta A. Sequence analysis and molecular characterization of the temperate lactococcal bacteriophage r1t. Molecular Microbiology 1996;19:1343–1355. [PubMed: 8730875]

Vezzi A, Campanaro S, D'Angelo M, Simonato F, Vitulo N, Lauro F, Cestaro A, Malacrida G, Simionati B, Cannata N, Bartlett D, Valle G. Genome analysis of Photobacterium profundum reveals the complexity of high pressure adaptations. 2004(GenBank Accession Number: NC_006370)

Wright, A.; McConnell, M.; Kanegasaki, S. Lipopolysaccharide as a bacteriophage receptor. In: Randall, LL.; Philipson, L., editors. Virus Receptors, Series B. Vol. 7. Chapman and Hall; New York: 1980. p. 27-58.

Young R. Bacteriophage lysis: mechanism and regulation. Microbiological Reviews 1992;56:430–481. [PubMed: 1406491]

Young R, Bläsi U. Holins: form and function in bacteriophage lysis. FEMS Microbiology Reviews 1995;17:191–205. [PubMed: 7669346]

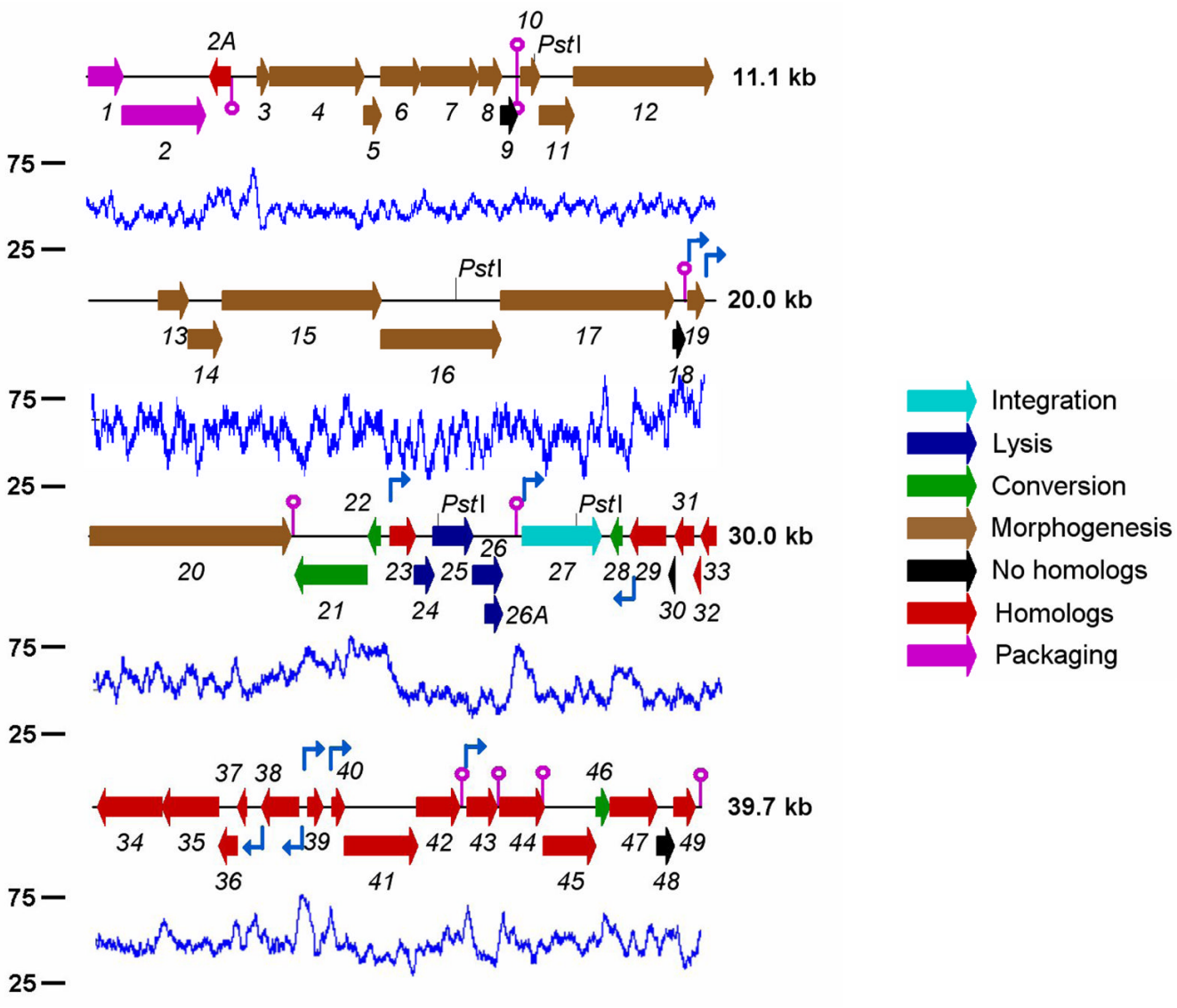

Fig. 1.

Gene diagram for *Salmonella* enterica, Serovar Anatum phage, ε15. Immediate and delayed early genes are designated in red, whereas the late genes are designated in brown. In addition, several genes and DNA regulator sites with known functions are depicted with unique colors, those being serotype conversion genes (green), the integrase gene (turquoise blue), transcriptional promoters (dark blue arrows), transcriptional termination signals (violet lollipops), lysis genes (dark blue), terminase genes (violet) and genes with no homologs in the GenBank (black). Also depicted are a 100 basepair running window of the AT%, as well as the positions at which PstI cuts the genome.

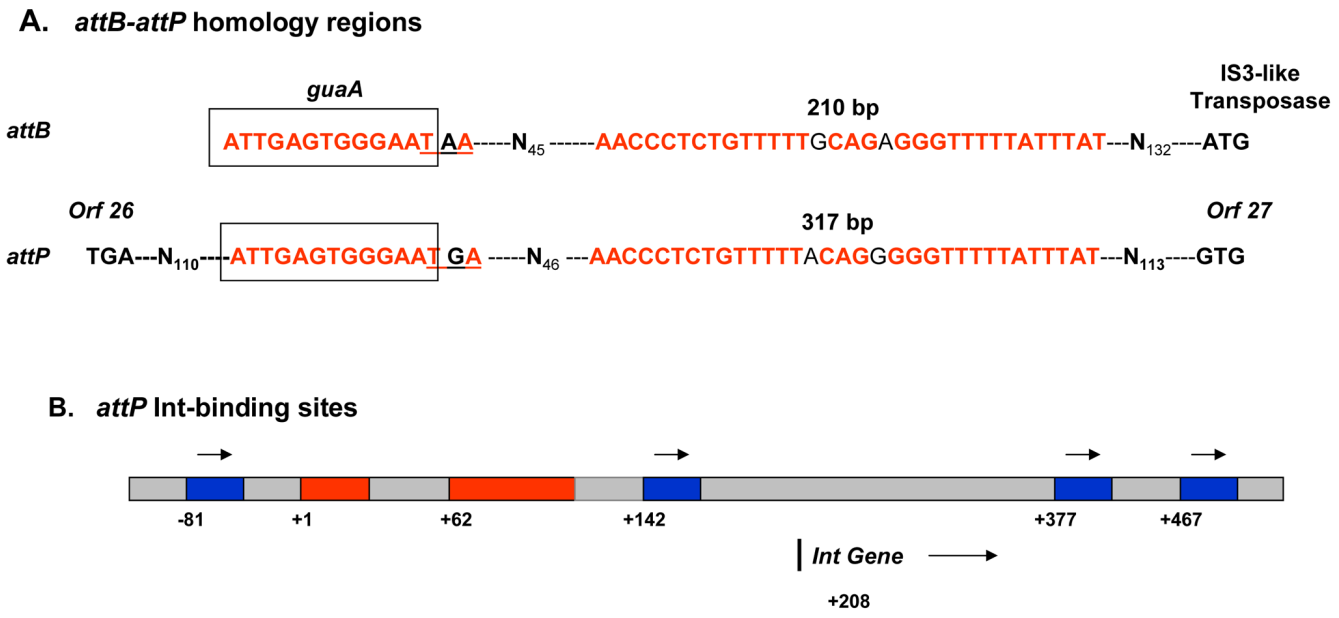

Fig. 2.

The attP/attB homology regions of ε15 and S. enterica, serovar Anatum.

Red font denotes the two common core regions of the attP and attB sites of E15 and *S. enterica*, Serovar anatum. Recombination occurs within the smaller, boxed common core region to the left.

The schematic diagram at the bottom of the Figure depicts the positions of the four likely "arm-type" binding sites for integrase (blue font), relative to the positions of the two "common core" sequences. The +1 designation refers to the first base of the smaller common core region wherein recombination occurs. The "arm-type" binding sites are all direct repeats of the sequence **GTGACGGTAT.** Also present, but not depicted, are two likely IHF binding sites (consensus **AATCAA**NNNN**TTR;** Goodrich, et al, 1990). One IHF binding site extends from -28 to -16 (**A**T**TCAA**TAAG**TTA**) and a second (**A**C**TCAA**TTAT**TTA**) is positioned within the complementary strand at a position (+7 through −6) that overlaps the smaller, common core sequence in which recombination occurs

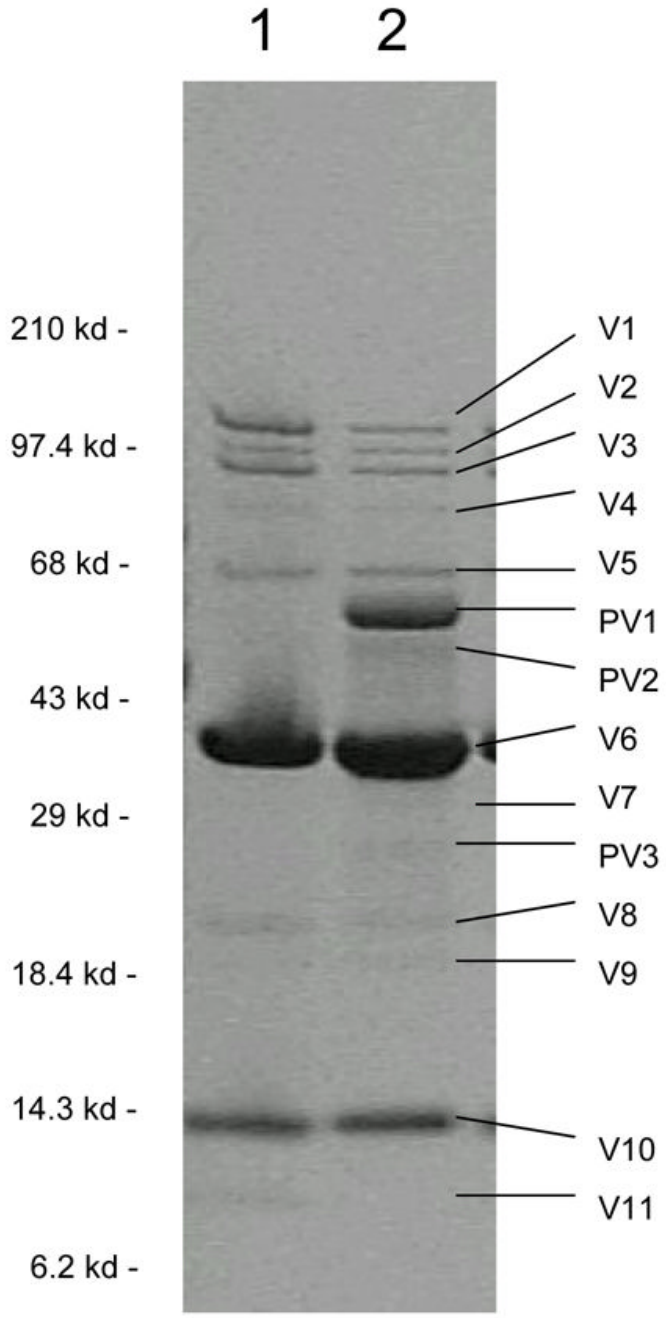

Fig. 3.

Polypeptide compositions of purified ε15 virion and 173S virion-like particles. The proteins of purified particles were resolved by SDS/Tricine/Polyacrylamide gel electrophoresis and visualized by CBB staining. Virion proteins are in Lane 1 and 173S particle proteins are in Lane 2. Each lane received a total of 23 ug of protein.

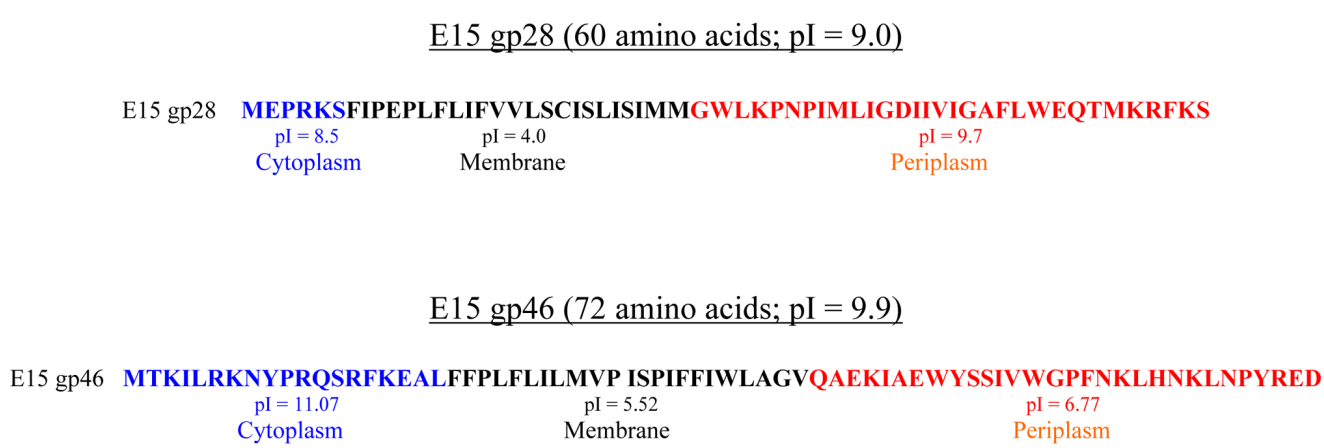

Fig. 4.

Membrane Topology of ε15 Cell Surface Conversion Proteins gp28 and gp46. According to the TMHMM program (Sonnhammer, et al., 1998), proteins gp28 and gp46 both have cytoplasmic N-terminal domains, a single membrane-spanning region and periplasmic C-terminal domains.

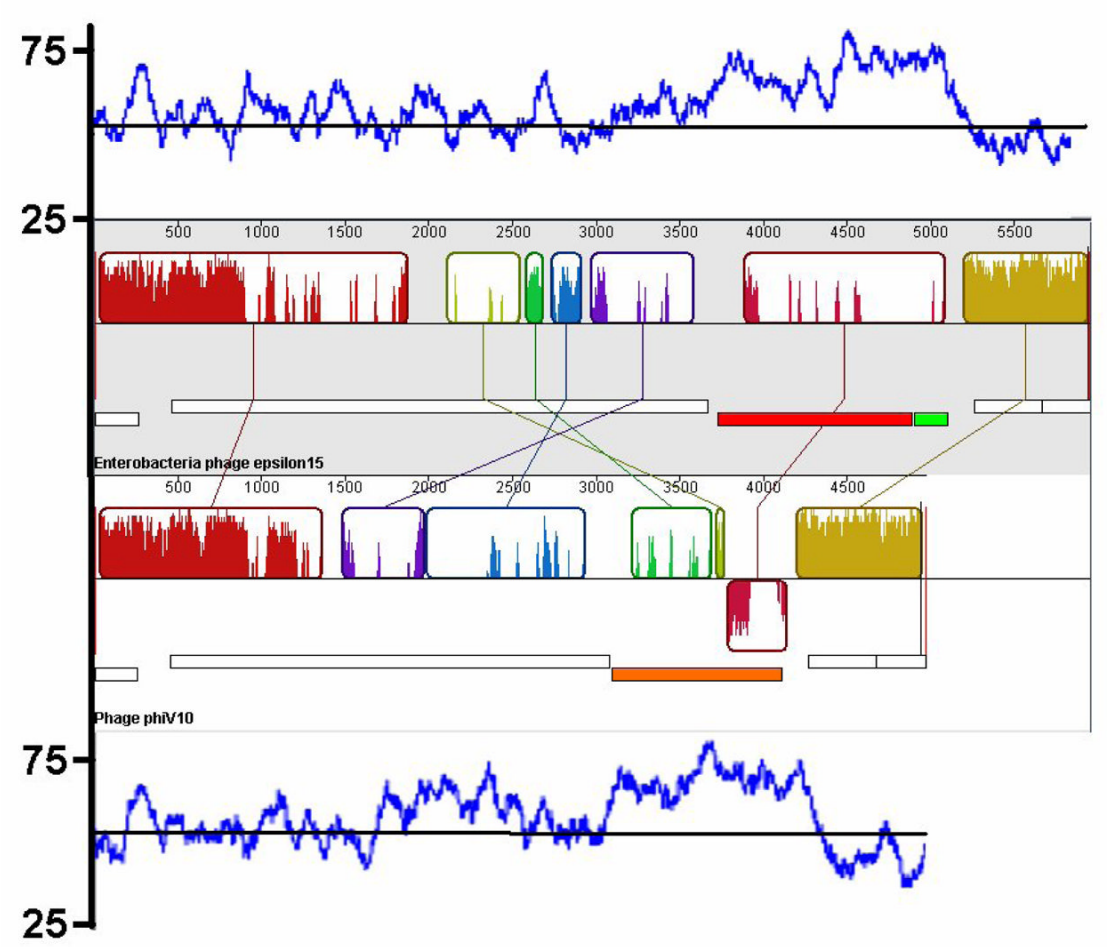

Fig. 5.

Comparison of the conversion modules of ε15 and ϕV10 using Mauve (Darling, et al., 2004). The horizontal boxes represent ε15 genes *19 through 24* and ϕV10 genes *24 through 28*, the longest of each, respectively, corresponding to the tailspike protein. The two ε15 genes involved in seroconversion are in red and green, while the putative ϕV10 transacetylase gene is in orange. The vertical colored blocks indicate regions of sequence similarity, the greater the degree, the greater the height of the colored bars. Above (ε15) and below (ϕV10) are genomic comparisons based upon scans of the AT-content of the respective regions.

Table 1

Coordinates of the genes of phage [15] with the properties of the protein products, related proteins and their functions, if known

	Position/Orientation					
ORF	**Begin**	**End**	**Mass/pI**	**Motifs**	**Putative function/Sequence similarity (gene & protein)**	**% Ident ity**
1	47	643 +	22695, 8.8	pfam03592	**Terminase small subunit** - *Escherichia coli* phage ϕV10 gene=PhiV10p01 [YP_512255]	77.3
2	640	2115 +	55117, 5.4	Prosite: PS00215	**Terminase large subunit** - *E. coli* phage ϕV10 subunit gene=PhiV10p02 [YP_512256]	94.3
2A	2181	2546 −	13271,4.9		- *E. coli* phage ϕV10 gene=PhiV10p03 [YP_512257]	78.0
3	3037	3243 +	6954, 6.5	-	- *E. coli* phage ϕV10 gene=PhiV10p06 [YP_512260]	88.2
4	3258	4928 +	61657, 4.9	-	**Putative head-tail connector protein (Portal)** - *E. coli* phage ϕV10 gene=PhiV10p07 [YP_512261]	79.6
5	4925	5221 +	11366, 4.9	-	- *E. coli* phage ϕV10 gene=PhiV10p08 [YP_512262]	64.3
6	5224	5940 +	25740, 4.6	-	**Endoprotease (putative)** - *E. coli* phage ϕV10 gene=PhiV10p09 [YP_512263]	69.7
7	5951	6958 +	36816, 6.2		**Major capsid protein** - *Bordetella* phage BIP-1 gene=BIP-1p16 [NP_996627]	46.0
8	6971	7363 +	13585, 4.5	-	- *Paracoccus denitrificans* gene=PdenDRAFT_1645 [ZP_00631364]	49.7
9	7356	7640 +	10686, 5.2	-	- *E. coli* phage ϕV10 gene=PhiV10p12 [YP_512266]	36.9
10	7705	8040 +	12167, 4.2	-	- *E. coli* phage ϕV10 gene=PhiV10p13 [YP_512267]	40.5
11	8040	8645 +	22204, 4.7	-	- *E. coli* phage ϕV10 gene=PhiV10p14 [YP_512268]	84.6
12	8645	11122 +	90871, 5.1	-	- *E. coli* phage ϕV10 gene=PhiV10p15 [YP_512269]	88.4
13	11122	11586 +	17392, 8.0	1 TMD	- *E. coli* phage ϕV10 gene=PhiV10p16 [YP_512270]	62.3
14	11586	12128 +	18354, 8.1	-	- *E. coli* phage ϕV10 gene=PhiV10p17 [YP_512271]	76.4
15	12141	14669 +	91011, 5.4	-	**Virion protein** - *E. coli* phage ϕV10 gene=PhiV10p18 [YP_512272]	57.4
16	14669	16573 +	67362, 4.4	-	**Virion protein** - *E. coli* phage ϕV10 gene=PhiV10p19 [YP_512273]	25.0
17	16573	19329 +	100840, 8.6	Prosite: PS00225	**Virion protein** - *S. glossinidius* gene=SG1195 [YP_454875]	27.2
18	19326	19520 +	7009, 8.9	-	-	
19	19559	19819 −	10009, 5.2	-	- *E. coli* phage ϕV10 gene=PhiV10p24 [YP_512278]	80.2
20	20017	23229 +	115616, 4.9	-	**Tailspike protein** - *E. coli* phage ϕV10 gene=PhiV10p25 [YP_512279] (N-Term)	25.4
21	23288	24448 −	43338, 9.1	10 TMD	β-**Polymerase involved in serotype conversion** - *Clostridium perfringens* gene=CPE0620[NP_561536)	25.5

	Position/Orientation					
ORF	**Begin**	**End**	**Mass/pI**	**Motifs**	**Putative function/Sequence similarity (gene & protein)**	**% Ident ity**
22	24461	24661 −	7515, 8.2	2 TMD	**α-Polymerase inhibitor involved in serotype conversion**	
23	24823	25227 +	14145, 5.8	3 TMD	- *E. coli* phage ϕV10 gene=PhiV10p27 [YP_512281]	92.5
24	25214	25522 +	11300, 11.0	3 TMD; pfam05449	**Holin** - *E. coli* phage ϕV10 gene=PhiV10p28 [YP_512282]	91.2
25	25512	26141 +	23211, 9.6	pfam00182 COG3179	**Endolysin** - *E. coli* phage ϕV10 gene=PhiV10p29 [YP_512283]	82.8
26	26138	26620 +	17854, 8.7	1 TMD	**Rz homolog (putative)** - *E. coli* phage ϕV10 gene=PhiV10p30 [YP_512284]	86.2
26A	26337	26612 +	9633, 8.4	1 TMD; PS00013 - lipoprotein	**Rz1 homolog** - *E. coli* phage ϕV10 gene=PhiV10p30a [YP_512285]	85.7
27	26938	28188 +	46780, 9.4	pfam00589	**Integrase** - *E. coli* phage ϕV10 gene=PhiV10p31 [YP_512286]	86.5
28	28343	28525 −	6940, 9.0	1 TMD	**Acetyltransferase inhibitor involved in serotype conversion**	68.1
29	28649	29227 −	21523, 7.7	Prosite: PS00092; pfam05063 COG4725	**Adenine Methylase** - *E. coli* phage ϕV10 gene=PhiV10p33 [YP_512288]	
30	29274	29381 −	4049, 5.4	-	-	
31	29374	29673 −	11243, 9.9	-	- *E. coli* phage ϕV10 gene=PhiV10p34 [YP_512289]	65.7
32	29670	29786 −	4684, 9.8	1 TMD	- *E. coli* phage ϕV10 gene=PhiV10p35 [YP_512290]	81.6
33	29783	30031 −	9414, 9.1	pfam05930; COG3311	**Transcriptional regulator** - *E. coli* phage ϕV10 gene=PhiV10p36 [YP_512291]	85.4
34	30080	31102 −	37287, 6.0	pfam03837	**RecT** - *S. glossinidius* gene=SG1175 [YP_454855]	72.9
35	31112	32011 −	33958, 6.1	pfam03837	**Endonuclease** - *S. glossinidius* gene=SG1176 [YP_454856]	65.6
36	32008	32310 −	11407, 4.6	-	-*E. coli* phage ϕV10 gene=PhiV10p38 [YP_512293]	51.0
37	32315	32464 −	5636, 8.6	-	- *E. coli* phage ϕV10 gene=PhiV10p39 [YP_512294]	76.0
38	32689	33285 −	22431, 8.1	smart00530; pfam01381	**Repressor (putative)** - *E. coli* phage ϕV10 gene=PhiV10p40 [YP_512295]	41.5
39	33441	33674 +	9139, 9.4	-	- *E. coli* phage ϕV10 gene=PhiV10p41 [YP_512296]	74.0
40	33819	34022 +	7770, 10.0	-	- *E. coli* phage ϕV10 gene=PhiV10p42 [YP_512297]	39.7
41	34019	35188 +	42402, 5.9	pfam00145; COG0270	**Cytosine Methylase** - *Pseudomonas putida* gene=PP1541 [NP_743698)	51.4
42	35178	35855 +	25970, 7.2	COG5529	- *S. glossinidius* gene=SG1211 [YP_454891]	23.2
43	35986	36450 +	16088, 9.3	pfam02075	**Crossover junction endodeoxyribonuclease RuvC homolog** - Candidatus Pelagibacter gene=PU1002_03366 [ZP_01264237]	34.9
44	36507	37202 +	25325, 4.4	pfam00607	**EaE homolog**	53.7

	Position/Orientation					
ORF	**Begin**	**End**	**Mass/pI**	**Motifs**	**Putative function/Sequence similarity (gene & protein)**	**% Ident ity**
					- *Salmonella* phage P22 gene=eae [NP_059592] (C- Term)	
45	37199	38032 +	30713, 4.9		**EaD homolog** - *Salmonella* phage P22 gene=ead [YP_063721]	66.8
46	38034	38252 +	8658, 9.9	1 TMD	**Acetyltransferase inhibitor involved in serotype conversion** - *Salmonella* phage ES18 gene=39 [YP_224177]	88.9
47	38256	39011 +	27780, 4.4	-	- *Salmonella* phage ES18 gene=38 [YP_224176]	31.1
48	39011	39283 +	10687, 10.2	pfam00170(leucine zipper)	-	
49	39276	39614 +	12506, 9.1	-	-*E. coli* phage ϕV10 gene=PhiV10p55 [YP_512310]	77.7

+/− strand containing the CDS; TMD = transmembrane domain

Table 2

Stoichiometry and Physical Characteristics of Proteins Contained in E15 Virions and 173S Particles

Protein Band #	Measured Protein Size[1] (and pI)[2]	Probable Encoding Gene With Predicted Size (and pI) of Inferred Protein Product[3,4]	Virion Quantity[5] Relative to Capsid	173S Quantity[5] Relative to Capsid
V1	115,508 ± 2,626 (nd)	Gene 20/115,616 (4.9) [3,4]	17.2	8.2
V2	94,923 ± 720 (nd)	Gene 17/100,840 (8.6)[4]	7.5	7.4
V3	88,053 ± 585 (5.74 ± 0.12)	Gene 15/91,011 (5.4)[3,4]	20.0	13.5
V4	74,249 ± 661 (4.58 ± 0.18)	Gene 16?/67,362 (4.4)[3]	9.6	4.0
V5	61,299 ± 528 (5.28 ± 0.13)	Gene 4/61,657 (4.9)[3,4]	20.0[6]	18.4
PV1	55,054 ± 1,628 (5.19 ± 0.08)	Gene 2/55,117 (5.4)[3]	-----	129
PV2	50,143 ± 1,209 (nd)	----------	-----	9.5
V6	35,192 ± 293 (6.38 ± 0.36)	Gene 7/36,816 (6.2)[3,4]	**415**	**415**
V7	29,734 ± 246 (nd)	----------	27.7	n.d.
PV3	26,978 ± 192 (4.61 ± 0.24)	Gene 6?/25,740 (4.6)[3]	-----	17.1
V8	21,910 ± 154 (nd)	----------	146	17.0
V9	19,273 ± 201 (nd)	----------	19.7	14.6
V10	12,087 ± 262 (4.81 ± 0.19)	----------	635	443
V11	7,826 ± 170 (nd)	----------	80.4	n.d.

[1] protein size estimates are the means (plus or minus the standard deviations) for six or more independent measurements

[2] protein pI values are the means (plus or minus the standard deviations) for four or more independent measurements

[3] gene identification supported by data presented in this paper

[4] gene identification supported by tryptic digestion and mass spectrometry data reported elsewhere (Jiang, Chang, et al, 2006)

[5] estimated number of polypeptide chains per average virion or 173S particle, calculated on the basis of CBB-stained band densities (n = 9) and measured polypeptide sizes, then normalized, assuming 415 capsid proteins per virion or 173S particle

[6] strong evidence (Jiang, Chang, et al, 2006) indicates that V5 (gp4) is the portal protein and that it is present at 12 copies per virion, suggesting that CBB may bind to V5 in a non-stoichiometric manner

Table 3

Identification of Cell Surface Conversion Genes Based Upon Phenotypic Characteristics of Transformed

Bacterial Strain[1]	Phage Sensitivity Pattern[2]				Whole Cell ELISA Results with Antibody to Antigen:		Appearance of LPS[3] on SDS-PAGE
	ε15	ε34	g341	Felix O1	O10 (alpha-linkages)	O15 (beta linkages)	
S. enterica A1 (wt)	S	R	S	R	+++	-	Smooth
S. enterica A1(ε15 lysogen)	R	S	R	R	-	+++	Smooth
S. enterica SR2 (wzy-)	R	R	R	S	-	-	Rough
S. enterica SR2/Orf *21*	R	S	R	R	-	++	Smooth
S. enterica A1/Orf *22*	R	R	R	S	-	-	Rough
S. enterica A1/Orf *28*	S	R	R	R	++	-	Smooth
S. enterica A1/Orf 46	S	R	R	R	++	-	Smooth

[1] S. enterica A1 (wt) is the strain used by Robbins and coworkers in their 1960s studies on serotype conversion by ε15. S. enterica SR2 (wzy-) is a mutant derivative of S. enterica A1 (wt) that lacks functional O-polysaccharide alpha polymerase enzyme and is therefore sensitive to “Rough”-specific bacteriophages (e.g. Felix O1)

[2] Phages ε15 and g341 both recognize alpha-linked Group E1 O-polysaccharide polymers; a difference is that phage g341 also requires that the galactose residues of the alpha-linked polymers be acetylated, whereas ε15 does not. Phage ε34 recognizes only beta-linked (Group E2) O-polysaccharide polymers. Felix 01 infects strains that either have no O-polysaccharide polymers, or else a single repeat unit of the O-polysaccharide.

[3] ”Smooth” denotes a ladder-like pattern of bands, indicative of a population of LPS molecules with varying numbers of repeat units in their O-polysaccharide polymers; “Rough” means that the ladder like pattern is absent and one sees instead one or two highly mobile bands, corresponding to complete Lipid A/R-cores with either zero or one O-polysaccharide repeat unit attached

Virology. Author manuscript; available in PMC 2009 June 18.