Author/s:
Dunstan, RA;Pickard, D;Dougan, S;Goulding, D;Cormie, C;Hardy, J;Li, F;Grinter, R;Harcourt, K;Yu, L;Song, J;Schreiber, F;Choudhary, J;Clare, S;Coulibaly, F;Strugnell, RA;Dougan, G;Lithgow, T

Title:
The flagellotropic bacteriophage YSD1 targets Salmonella Typhi with a Chi-like protein tail fibre

Date:
2019-12-01

Citation:
Dunstan, R. A., Pickard, D., Dougan, S., Goulding, D., Cormie, C., Hardy, J., Li, F., Grinter, R., Harcourt, K., Yu, L., Song, J., Schreiber, F., Choudhary, J., Clare, S., Coulibaly, F., Strugnell, R. A., Dougan, G. & Lithgow, T. (2019). The flagellotropic bacteriophage YSD1 targets Salmonella Typhi with a Chi-like protein tail fibre. Molecular Microbiology, 112 (6), pp.1831-1846. https://doi.org/10.1111/mmi.14396.

DR. RHYS GRINTER (Orcid ID : 0000-0002-8195-5348)
PROF. RICHARD A STRUGNELL (Orcid ID : 0000-0003-0614-5641)
DR. TREVOR LITHGOW (Orcid ID : 0000-0002-0102-7884)

Article type : Research Article

The flagellotropic bacteriophage YSD1 targets *Salmonella* Typhi with a Chi-like protein tail-fibre.

Running title: Flagella targeting by bacteriophage YSD1

Rhys A. Dunstan[1,+], Derek Pickard[2,3,+], Sam Dougan[2], David Goulding[2], Claire Cormie[2], Joshua Hardy[4], Fuyi Li[4], Rhys Grinter[1,5], Katherine Harcourt[2], Lu Yu[2], Jiangning Song[4], Fernanda Schreiber[2], Jyoti Choudhary[2], Simon Clare[2], Fasseli Coulibaly[4], Richard A. Strugnell[6], Gordon Dougan[2,3,*] and Trevor Lithgow[1,*]

1 Infection and Immunity Program, Biomedicine Discovery Institute and Department of Microbiology, Monash University, Clayton 3800, Australia
2 Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, United Kingdom
3 Department of Medicine, University of Cambridge, Addenbrooke's Hospital, Hills Road, Cambridge, UK.
4 Infection and Immunity Program, Biomedicine Discovery Institute and Department of Biochemistry & Molecular Biology, Monash University, Clayton 3800, Australia
5 School of Biological Sciences, Monash University, Clayton 3800, Australia
6 Department of Microbiology and Immunology, The Peter Doherty Institute, The University of Melbourne, Parkville 3052, Australia

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/MMI.14396

Keywords: Bacteriophage/Flagella/natively disordered protein/Siphoviridae

+ these authors contributed equally
* e-mail: gd312@medschl.cam.ac.uk , trevor.lithgow@monash.edu

SUMMARY

The discovery of a *Salmonella*-targeting phage from the waterways of the United Kingdom provided an opportunity to address the mechanism by which Chi-like bacteriophage (phage) engage with bacterial flagellae. The long tail fibre seen on Chi-like phages has been proposed to assist the phage particle in docking to a host cell flagellum, but the identity of the protein that generates this fibre was unknown. We present the results from genome sequencing of this phage, YSD1, confirming its close relationship to the original Chi phage and suggesting candidate proteins to form the tail structure. Immunogold-labelling in electron micrographs revealed that YSD1_22 forms the main shaft of the tail-tube, while YSD1_25 forms the distal part contributing to the tail-spike complex. The long curling tail fibre is formed by the protein YSD1_29, and treatment of phage with the antibodies that bind YSD1_29 inhibit phage infection of *Salmonella*. The host range for YSD1 across *Salmonella* serovars is broad, but not comprehensive, being limited by antigenic features of the flagellin subunits that make up the *Salmonella* flagellum, with which YSD1_29 engages to initiate infection.

INTRODUCTION

Phages are viruses that infect bacteria. The best-studied order, the *Caudalovirales*, conform to a blue-print consisting of a protein capsid housing a double stranded DNA genome, joined by a portal structure to a multi-component tail. The function of the tail is to engage a receptor on the surface of the host bacterium, in order to initiate the infection process (Nobrega *et al.*, 2018). Conceptually, the tail is best considered as a molecular machine, as engagement with the host bacterium and subsequent injection of the genomic DNA requires intricate ligand-receptor interactions and movements of the distal and proximal components in the tail (Nobrega *et al.*, 2018). The *Caudalovirales* are further divided into three major families based on their tail morphology: the *Podoviridae* (short-tailed, non-contractile phage),

Myoviridae (long-tailed, contractile phage) and the *Siphoviridae* (long-tailed, non-contractile phage) (Ackermann & Prangishvili, 2012).

Phage Lambda is the archetypal *Siphoviridae* (Casjens & Hendrix, 2015). Lambda was isolated in 1951 (Lederberg, 1951) and its genome was later sequenced in 1982 (Sanger *et al.*, 1982). Lambda phage selectively infects *Escherichia coli* through short tail fibres that make specific interactions with the ß-barrel protein LamB in the bacterial outer membrane (Charbit, 2003, Chatterjee & Rothenberg, 2012). A large group of phages, referred to as the Lambda supercluster and presenting with mosaic genome-based relationships are now known, and have uniquely informed our understanding of how gene-swapping drives phage evolution (Grose & Casjens, 2014). In 1936, a distinct phage of the class *Siphoviridae* was discovered that could infect *Salmonella enterica* and which were called Chi phage (Sertic & Boulgakov, 1936, Meynell, 1961). Chi phages have a single long tail fibre, distinguishing them morphologically from Lambda and, uniquely, Chi phages are reported to bind to the flagellae on *Salmonella* cells (Schade *et al.*, 1967, Ravid & Eisenbach, 1983, Samuel *et al.*, 1999). Recently, the genome of bacteriophage Chi was sequenced and the 59,578 base-pair sequence was predicted to encode 75 proteins, most of these having no known function (Hendrix *et al.*, 2015).

Chi phage serves as the prime example of a group of closely related Chi-like phages (Moreno Switt *et al.*, 2013, Grose & Casjens, 2014, Hendrix *et al.*, 2015), most of which infect *Salmonella enterica* serovars, but some of which are reported to be specific for *Enterobacter cancerogenous* (Kazaks *et al.*, 2012) or *Providencia stuartii* (Onmus-Leone *et al.*, 2013). Diverse flagellotropic phages have been identified to infect a range of bacterial species including *Aeromonas*-phage PM3 (Merino *et al.*, 1990), *Agrobacterium*-phages 7-7-1, GS2 and GS6 (Kropinski *et al.*, 2012, Bradley *et al.*, 1984), *Asticcacaulis*-phages and *Caulobacter*-phages ϕAcS2, ϕAcM4, ϕCp34, ϕCb13 and ϕCbK, and ϕ6; (Fukuda *et al.*, 1976, Guerrero-Ferreira *et al.*, 2011, Jollick & Wright, 1974, Pate *et al.*, 1979), *Bacillus*-phages AR9, 3NT, PBS1, SP3 and PBP1 (Vieira *et al.*, 1989, Shea & Seaman, 1984, Lovett, 1972), *Campylobacter*-phage F431 (Baldvinsson *et al.*, 2014) and *Proteus*-phage PV22 (Zhilenkov *et al.*, 2006), *Pseudomonas*-phage ϕCTX (Geiben-Lynn *et al.*, 2001). Interestingly, for ϕCb13 and ϕCbK, targeting the flagellum of their host is performed by a flexible filament emanating from its capsid rather than the distal tail complex observed in Chi phage for example (Guerrero-Ferreira *et al.*, 2011).

While fewer in number to date, other phages have been discovered to use pili as their primary receptor (Harvey *et al.*, 2018, McCutcheon *et al.*, 2018). Kropinski *et al.* (2012) have argued that this general mechanism of binding to extended filaments increases the probability of encountering the target bacterium, since bacteria with their flagella extended have a capture radius 5- to 10-fold greater than the bacterial cell itself (Kropinski *et al.*, 2012). Furthermore, once the encounter with the flagellum has been made, estimates suggest that Chi-phage needs less than a second to reach the bacterial cell surface (Samuel *et al.*, 1999). While these extended receptor structures enhance the capture radius for phage-engagement, recent evidence suggests the importance of a secondary receptor at the outer membrane surface to initiate genome ejection into the bacterial cell. In the case of *Agrobacterium*-phage 7-7-1 and two Chi-like phages, STm101 and STm118, lipopolysaccharide (LPS) has been shown to be important in the infection process of these phages (Gonzalez *et al.*, 2018, Phothaworn *et al.*, 2019).

The host range of Chi phage within the Salmonellae species was found to have a consistent pattern dependent largely upon the flagella serotype (Meynell, 1961). Recently other *Salmonella* flagella targeting phages have been described (Shin *et al.*, 2012, Moreno Switt *et al.*, 2013, Phothaworn *et al.*, 2019) and all share a number of genes that can be used as markers for Chi-like phages. We present the results from genome sequencing of the phage, YSD1, which we isolated from an environmental survey of water sources in the United Kingdom. Sequencing revealed its close relationship to the original Chi phage, placing it in the Chi-like phage cluster. The 58,916 base-pair genome of YSD1 has the same packaging ends as Chi, and is predicted to encode 71 proteins. Mass spectrometry showed that at least 24 of these proteins are found in the mature virion. Gene synteny with the archetype Chi phage was extensive, with only three genes in YSD1 not shared by Chi; none of these three gene products were found in the YSD1 virion. Electron microscopy showed that the tail tube of YSD1 is formed from at least two proteins: YSD1_22 forms the main shaft of the tail-tube, while YSD1_25 forms the distal part contributing to the tail-spike. With the long flexible tail fibre protein being formed by YSD1_29. Given the 99% sequence identity with the homologous structural proteins in Chi, the results on YSD1 serve as a paradigm for the Chi-like group of phages and provide a vehicle to study the interaction of flagella with the phage and the role of the long tail fibre in the process.

RESULTS

Isolation and characterisation of YSD1

The phage YSD1 was isolated from environmental water samples taken during a phage survey of the River Cam, Cambridge UK, using the attenuated *S. enterica* serovar Typhi BRD948 as the bacterial host. When grown in the presence of *S.* Typhi BRD948, YSD1 showed lytic activity producing small, clear plaques of 0.5 to 1 mm in size (Fig. S1A). Some phage use the Vi capsule of *S.* Typhi as a receptor (Pickard *et al.*, 2008), but when tested against an *S.* Typhi BRD948 derivative that does not express the Vi capsule, YSD1 still infected the bacteria (Fig. S1B). Transmission electron microscopy (TEM) of purified phage samples (Fig. 1A, B) revealed that the YSD1 virion has a morphology consistent with that of a phage belonging to the *Siphoviridae* family; with an icosahedral capsid (~60 nm in diameter), noncontractile tail (~220 nm in length). A long, curling tail fibre situated at the end of the major tail segment (Fig. 1A, inset) suggested it may belong to the Chi-like group of phages.

When co-cultured with *S.* Typhi BRD948, TEM analysis showed that the long tail fibre on YSD1 interact with the bacterial flagellum (Fig. 1C). YSD1 was often observed at the base of the flagellae (Fig. 1D) consistent with it migrating along the flagellum to the surface of its bacterial host. Whether captured amidst the flagellar filament or at its base, the YSD1 virions had capsids that were turgid (Fig 1D, arrowheads), suggesting that they had yet to release their DNA contents. The bacterial flagellum is composed of flagellin subunits, with the FliC flagellin being common to *Salmonella* spp. and other species of Enterobacteriaceae. YSD1 did not infect *S.* Typhi if the *fliC* gene was deleted (Fig. S1A). A tree depicting the sequence conservation in FliC mapped out the FliC proteins from selected *Salmonella* strains susceptible to infection with YSD1 (Fig. 1E).

Genome characteristics of YSD1

Sequencing of the YSD1 genome revealed it to be a single, linear 58,916 base pairs polynucleotide with a G+C content of 56.63% (Genbank accession: LR026998). The genome consists of 71 predicted open-reading frames (Table 1), no tRNA genes, and has the same terminal 12 base-pair sequence: GGTGCGCAGAGC, found at both ends of the Chi phage genome (Genbank accession: KM458633). When the YSD1 genome was ordered based upon these cohesive ends, a pair-wise sequence analysis showed YSD1 and Chi shared ~ 97%

sequence identity in nucleotide sequence (Fig. 2A) that can be organized in terms of functional regions (Fig. 2B). At the protein sequence level, the predicted structural genes are 99-100% identical (Table 2). Thus, at a protein structural level, YSD1 and Chi virions are indistinguishable. At the genome level, small non-homologous region distinguishes YSD1 from Chi (Fig. 2A, inset): two genes (encoding gp42 and gp48) are present in Chi but absent from YSD1. Instead, YSD1 has a gene that encodes a distinguishing protein referred to as YSD1_40 (Fig. 2A). Nothing is known about the function of this protein, though the gene encoding YSD1_40 has been identified in three other Chi-like phages: iEPS5 (NCBI accession: NC_021783) and SPN19 (NCBI accession: NC_019417), as well as the Enterobacter phage Enc34 (GenBank accession: JQ340774). Conversely the gene encoding the hypothetical protein gp42 from Chi phage, absent in the genome of YSD1, is also present in the genome of other Chi-like phages including iEPS5. While a relatively minor example, this genome mosaicism is reminiscent of the features described in the Lambda-like superfamily of phages (Grose & Casjens, 2014). This very small distinction in what is otherwise an identical phage suggests YSD1 and Chi diverged from a common ancestor relatively recently.

Architecture of proteins in the tail structure of YSD1

YSD1 was a useful model to identify and determine the candidate structural proteins that comprise the architectural features in the tail-tube, tail-spike and curling tail fibre for the Chi-like family. Table 1 documents outcomes from analyses with BLASTp (https://blast.ncbi.nlm.nih.gov/Blast.cgi), HMMer [(Finn *et al.*, 2011, Potter *et al.*, 2018); https://www.ebi.ac.uk/Tools/hmmer/], HHpred [(Zimmermann *et al.*, 2018); https://toolkit.tuebingen.mpg.de/#/tools/hhpred] and Phyre2 [(Kelley *et al.*, 2015); http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index] searching for domains generally characteristic of phage tail proteins, with candidates: YSD1_22, YSD1_25 and YSD1_29, and also putative tail assembly chaperones (YSD1_21, YSD1_23, YSD1_26, YSD1_27 and YSD1_28), as well as a tape measure protein (YSD1_24). When assessed by SDS-PAGE, the protein profile for the virions of YSD1 showed two highly-abundant proteins (the putative major capsid and tail proteins) present in hundreds of copies per virion, in addition to other proteins present at reduced stoichiometry (Fig. 3A). To independently address which proteins would be abundant as structural components of the phage virions, mass spectrometry was used and confirmed the presence of 24 protein species in the virions, several of which are so far unannotated (Table 2).

A predicted tail protein, YSD1_22, carries a C-terminal bacterial Ig-like 1 domain and is 99% identical to protein gp22 from Chi phage. Automated annotation based on sequence characteristics suggested YSD1_22/gp22 as a "major tail protein". YSD1_25 is predicted as a "distal tail protein" showing 99 % sequence identity to gp26 from Chi phage. YSD1_29 contains a predicted Phage-tail_3 domain, and is 99 % identical to the "minor tail protein" gp30 from Chi phage. All three of these proteins were identified by mass spectrometry as being present in the virion (Table 2). To experimentally address the localisation of these three proteins in the YSD1 virion, C-terminally his-tagged YSD1_22, YSD1_25 and YSD1_29 were expressed in recombinant form and purified for polyclonal antibody production.

Immunoblot analysis demonstrated the specificity of each antibody, revealing a unique band at the appropriate size (Fig. 3B), and immunogold-labelling of YSD1 using these antibodies labelled specific regions of the phage. YSD1_22 is located along the shaft of the tail-tube (Fig. 3C). Allowing for the length of the antibody-antibody-gold detection reagent, YSD1_25 is located in close proximity to the tail-spike complex, distal to the head (Fig. 3D). While localized to the same region of the phage virion, the antibody-antibody-gold detection reagent recognizing YSD1_29 showed an even more dispersed labelling (Fig. S2), consistent with it representing the position of the long, curling tail fibre (Fig. 3E). YSD1_29 is a 143 kDa protein (Table 2, Fig. S3A) with several independent software tools suggested that the YSD1_29 protein is natively disordered over some of its length (Fig. 3F), and HMMER predicting a Phage-tail_3 domain spanning residues 208-380. To probe the structural features of the protein further, purified YSD1_29 was analysed by size-exclusion chromatography and small-angle X-ray scattering (SAXS). YSD1_29 proved to be sensitive to proteolysis during purification, with a segment corresponding to 372 residues from the N-terminus degraded (Fig. S3A). Despite this, the ~105 kDa protein that remained (YSD1_29$^{373\text{-}1296}$) was stable, and was purified for analysis by SAXS (Table S1). In solution, the protein was observed to be flexible (Fig. S3C), consistent with the disorder predictions, and to be elongate in shape (Fig. S3D). Overall, the SAXS analysis revealed a cork-screw shaped protein with a cross-sectional diameter of ~4-5 nm. The fibre-like structure of monomeric YSD1_29$^{373\text{-}1296}$ has a calculated length of 20.9 nm (Fig. 3G).

Characterisation of YSD1-flagella interactions

S. enterica serovars can switch flagellae to various forms by altering the expression of flagellins. In the well-documented case of *S. enterica* serovar Typhimurium (Fig. 4A), a genetic switch enforces expression of either FliC or FljB for assembly into the flagella filament (Bonifield & Hughes, 2003). To define infection parameters for an assay to measure flagellae-dependent binding of YSD1, we made use of this defined bacterial system as a host for phage infection. A stock of YSD1 (10^{10} PFU ml^{-1}) was serially diluted from 10^{-1} to 10^{-6} and volumes of 10 µl were spotted onto a lawn of *S*. Typhimurium or the *ΔfliC,ΔfljB* mutant of *S*. Typhimurium. Visible plaques were observed up to 10^{-6} on the wild-type and 10^{-2} on the *ΔfliC,ΔfljB* mutant, providing the range of infection parameters over which flagella-dependent binding can be monitored (Fig. 4B). That the *ΔfliC,ΔfljB* mutant is sensitive to YSD1, albeit at very high MOI values, is consistent with the presence of a secondary receptor as has been suggested for the flagellotropic phages 7-7-1, STm101 and STm118 (Gonzalez *et al.*, 2018, Phothaworn *et al.*, 2019).

We sought to test whether the tail fibre protein YSD1_29 mediates the binding to the primary receptor, the flagellum. Antibodies recognizing YSD1_29 were pre-adsorbed onto YSD1 particles prior to mixing with bacteria under these defined conditions (Table 3). Serial dilutions of the YSD1 stock described above (10^{-1} to 10^{-6}) and the antibody against YSD1_29 (10^{-1} to 10^{-4}) were prepared, mixed in a 1:1 ratio and incubated at 37 °C for 1 hour. The phage-antibody mix was then spotted onto a lawn of *S*. Typhimurium SL1344 and incubated overnight at 37 °C. No visible plaques developed if the antibody to YSD1_29 was used at a dilution of 10^{-1}, regardless of phage concentration (Table 3, Fig. S4). This was in comparison to full zones of clearance in the cells spotted with phage alone (Table 3). There was no such inhibition seen if antibodies to YSD1_22 and YSD1_25 were used instead. Taken together with the localisation and structural studies, we suggest that YSD1_29 forms the flexible tail-fibre in Chi-like phages, and that it plays a substantive role in binding to bacterial flagellae.

Specific flagella antigens determine YSD1 host-range

In the serological typing system used to classify Salmonellae (Guibourdenche *et al.*, 2010), the flagellar filament provides the *H* antigen. This system therefore provides a means to map sequence variation in the flagellin subunits. To understand whether the YSD1_29 protein in YSD1 distinguished between these flagellae, a panel of thirty *Salmonella* serovars were tested for sensitivity to YSD1. To quantify and compare the outcomes, 10^{10} PFU mL^{-1} of

YSD1 were diluted (10^{-1} to 10^{-6}) and liquid infections were performed to define an approximate efficiency of plating (EOP) score, to determine for the various strains the degree of sensitivity to YSD1 (Methods). Approximate EOP values were calculated with reference to *S*. Typhi that expressed H*d* flagella and, using these assay conditions, no plaques were formed on the non-flagellated control *S. enterica* serovar Gallinarum biovar Pullorum. Under conditions of this assay, YSD1 formed plaques on most strains, except those that expressed either *h*, *g* or z_{66} antigenic form of flagellae (Table 4). A geographically restricted variant of the H*d* flagella, referred to as H*j*, was observed in *S*. Typhi strains isolated in Indonesia (Frankel *et al.*, 1989). In H*j* strains, the variant FliC protein isoform has a large in-frame deletion that removes 87 amino acids within the D2 and D3 domains of FliC (Fig. S5A). Strains that produce this H*j* variant still form functional flagella filaments (Schreiber *et al.*, 2015), and they are as sensitive to YSD1 as the wild-type H*d* variant of *S*. Typhi as calculated by EOP (Table 4, Fig. 4C).

A recent theoretical model to describe the movement of flagellotropic phage along flagellar filaments proposes that sequence specific determinants in the flagellin subunits would determine the speed and extent of translocation along the flagellum (Katsamba & Lauga, 2019). A fine-grained analysis of flagellin sequences from the various *Salmonella* serovars provided a framework on which to analyse the data from Table 4. As expected, this grouping segregates the serovars according to their *H* antigen designations (Fig. 4C). The tree reveals that the sensitive subgroups form discreet lineages. The z66/g lineage and the h lineage each constitute a group of serovars resistant to YSD1 (Fig. 4C). Fig. 4D illustrates the four domains of FliC and how they are organized in a bent hair-pin structure (Yonekura *et al.*, 2003). In the assembled flagellum, the domains of FliC comprising the flagella filament core are designated D0 and D1, and the D2 and D3 domains projected radially outwards from this core. Sequence conservation analysis across the serovars, mapped onto this structure, revealed that it is the D2 and D3 domains that are the most variable (Fig. 4E). Strings of amino acid homology within the D2/D3 domains were notable for many phage resistant hosts in comparison to those hosts sensitive to the phage.

In the context of an assembled flagellum filament, the D2 and D3 domains form a surface groove, the properties of which would be altered by sequence variation and sequence deletion (Fig. 4F). Variations in sequence for the D2 and D3 domains change the electrostatic features of the flagellum (Fig. S5B). These variations include a more positively-charged patch on the

surface of *S*. Typhi Hz66 and *S*. Enteritidis, which correlate with phage resistance. Past speculations that flagellotropic phage might inject the genomic DNA directly into the flagellum are not supported by the structural model: while a central channel is present, and is sufficiently large for a linear DNA molecule (Fig. 4G), there is no gap between the flagellin subunits (Fig. 4G) that would afford entry from the outside surface into this channel.

DISCUSSION

Phage are non-motile and encounter bacterial surfaces by chance. A specific, receptor-binding, interaction with a bacterial surface feature ensures that the infection process is not initiated until the phage has encountered its *bona fide* host. Only after that interaction has occurred does a cascade of events ensue to ultimately see the genome of the phage translocated through the tail-tube to enter the bacterial cytoplasm for expression of phage genes and replication of phage virions. In the case of phage Chi, the primary receptor that it interacts with is the bacterial flagellum. Here we investigated the molecular details of that interaction for the Chi-like phage, using phage YSD1 as a model.

Three proteins: YSD1_22, YSD1_25 and YSD1_29, are structural elements in the tail of YSD1. These proteins have near identical (99% sequence identity) counterparts in Chi and other Chi-like phages. YSD1_22 is associated with the tail tube, whereas YSD1_25 and YSD1_29 form the distal elements of the tail. Antibodies to YSD1_29 recognised the long, flexible tail-fibre. The protein has a protease-resistant domain that accounts for ~20 nm of the length of this fibre as judged by SAXS analysis. This flexible structure was observed in immune-stained electron micrographs of YSD1 (Fig. 3E), as well as in flagellar-located YSD1 (Fig. 1C) and Chi (Schade *et al.*, 1967). It corresponds to the main element that would need to interact with bacterial flagellae in current theoretical models for engagement of flagellotropic phage with bacteria (Katsamba & Lauga, 2019). *Salmonella* spp. are endemic microflora in birds, and a very recent survey of poultry farms in Thailand identified the Chi-like phages STm101 and STm118. In that survey, a PCR assay for amplification of the gene encoding capsid protein of the Chi-like phage showed that half of all phage isolates from across the farms corresponded to Chi-like phages (Phothaworn *et al.*, 2019).

Within the *S. enterica* serovars tested here, we found that some are resistant to infection with YSD1. The correlation between *H* antigen status (i.e. flagellin sequence features) and susceptibility to YSD1 infection is consistent with YSD1 being flagellotropic. Given the

ready prospect of evolving a flagellum that is phage-resistant, what is the advantage to a phage in using flagellae as a receptor? Previous ideas that the flagellum would form a conduit for the phage genome to enter the bacterial cytoplasm were attractive, but cryo-electron microscopy of the flagellum does not allow sufficient entry space into the filament channel and protein export (i.e. macromolecular transport out through the flagellum) would otherwise completely occupy the internal space (Minamino, 2014). This modelling, coupled with observations that the YSD1 observed on the flagellae have turgid capsids, suggest that the flagellum serves as a cable down which the phage ratchets to the bacterial surface to encounter a secondary receptor that triggers DNA release. The observation that, albeit at greatly lowered efficiency, the *ΔfliC*,*ΔfljB* mutant is infected by YSD1 further supports the contention that encounter with a secondary receptor in the bacterial outer membrane might be the trigger for genome release from the Chi-like phage capsid. Visible plaques were observed using spot assays with high titres of YSD1 on the *ΔfliC*,*ΔfljB* mutant that lacks flagellae (Fig. 4B), as has been observed with the *Agrobacterium*-phage 7-7-1 (Gonzalez *et al.*, 2018). In this case and similarly to that of Chi-like phages STm101 and STm118, it is LPS which serves the function of a secondary receptor on the bacterial cell surface (Gonzalez *et al.*, 2018, Phothaworn *et al.*, 2019).

Movement along the bacterial flagellum was first proposed as a nut-and-bolt mechanism (Berg & Anderson, 1973) and has recently been modelled through precise calculations that indicate that speed of travel is determined by factors that include the length of the tail fibre (Katsamba & Lauga, 2019). In *Salmonella*, counter-clockwise rotation operates for forward swimming, a state in which each flagellum forms a coordinated bundle (Macnab, 1977). Theoretically, clock-wise rotation of the flagellum, which enables bacteria to tumble and sample their environment, would force phage away from the bacterium (Katsamba & Lauga, 2019). However, given that *Salmonella* flagellae bundle together during counter-clockwise rotation (Macnab, 1977), this locomotive need is used against the bacterial host and ensures phage entrapment on the fibres to subsequently infect the bacterial cell swimming towards nutrients.

In YSD1, Chi and Chi-like phages, YSD1_29 and its homologs forms the tail fibre. The hydrodynamic calculations of phage movement along flagellae support a model in which only a counter-clockwise rotation of the flagellum would promote movement of the phage along the right-handed groove in the flagellum towards the bacterial cell surface (Katsamba &

Lauga, 2019). The corkscrew shape of the YSD1_29 protein adds further support to this intimate relationship between the phage tail structure and the bacterial flagellum. In addition to the rotation of the flagella filament, the targeting is also largely dependent on the putative groove properties formed by the variable D2 and D3 domains of the flagellin protein. In the case of YSD1 and Chi, their host range is broad because diverse Salmonella express conserved flagellae, but those serotypes of *Salmonella* expressing distinct flagellae are resistant to YSD1 as they would be to the near identical Chi virions. This study provides a better understanding of how Chi-like phage target the flagella of their target bacterial host and may provide insight into the targeting of other flagellated bacterial species by their respective phages.

EXPERIMENTAL PROCEDURES

Bacterial strains

The attenuated *Salmonella enterica* serovar Typhi strains BRD948 (Tacket *et al.*, 1997), *S.* Typhi BRD948 *ΔfliC* (Schreiber *et al.*, 2015) and *S.* Typhi strain BA256 [Vi capsule mutant *ΔtviB*; (Pickard *et al.*, 2010)] were used for YSD1 isolation and initial infections. All bacterial strains that were used to test YSD1 sensitivity are listed in Table S2. Protein expression strain *E. coli* C41 (Lucigen) was used for YSD1_22, YSD1_25 and YSD1_29 expression for antibody production and structural studies.

Phage Infections

Phage infections were either performed in liquid media or by spotting onto a bacterial lawn. For liquid infections 10 µL of serially diluted YSD1 was added to 200 µL of bacterial culture and incubated at 37 °C for 20 min to allow phage adsorption. Luria Broth (4 mL) containing 0.35% (w/v) noble agar was added to the phage-bacteria mix and poured using the double overlay method. For spot assays 4 mL of Luria Broth (0.35% (w/v) noble agar) was added to 200 µL of bacterial culture and poured as described above. Plates were allowed to set and 10 µL of serially diluted YSD1 was spotted onto the top agar. Plates were subsequently incubated at 37 °C overnight for plaque formation.

Efficiency of plating (EOP)

Efficiency of plating (EOP) was calculated by determining the approximate ratio of plaque-forming-units per millilitre (PFU mL^{-1}) produced by a YSD1 infected *Salmonella* serovar

compared to *S.* Typhi BRD948. An EOP of less than 1 would indicate strains less susceptible to YSD1 infection compared to *S.* Typhi BRD948.

YSD1 Genomic DNA preparation

YSD1 genomic DNA was prepared from a 1.8 mL sample of phage lysate (~10^{10} PFU ml^{-1}). Eighteen microlitres of DNase (1 mg mL^{-1}) and 8 µL RNase A (12.5 mg mL^{-1}) were added and incubated at 37 °C for 30 min. Eighteen microlitres of proteinase K (10 mg mL^{-1}) and 46 µL SDS (20% stock solution) were added to the lysate and subsequently incubated for a further 30 min at 37 °C. Samples (500 µL) were aliquoted to 1.5 mL phase-lock tubes for phenol chloroform extraction. Phenol:Chloroform:Isoamyl Alcohol (25:24:1) (500 µL) was added to each tube and centrifuged for 5 min at 1500 x g. The top aqueous layer was removed and added to a fresh phase-lock tube. Chloroform:Isoamyl Alcohol (24:1) (500 µL) was added and centrifuged for 5 min at 6000 x g and the aqueous layer was removed and aliquoted to a microcentrifuge tube. Forty-five microlitres of 3M Sodium Acetate pH5.2 and 500 µL of 100% Isopropanol were added to each tube. DNA was precipitated at room temperature for 20 min and subsequently centrifuged for 20 min at 14000 rpm. The DNA pellet was washed twice with 70% Ethanol prior to drying and resuspension in 100 µL of milliQ water for storage.

Comparative sequence analysis

The Flagellin sequence dataset for phylogenetic analysis was retrieved using functionally described sequences containing a 'Flagellin_N' pfam domain (PF00669; Table S3) from a curated list of bacterial species within the Enterobacteriaceae family (Pfam 32.0, http://pfam.xfam.org, retrieved 2018, Dec 12). A second Flagellin dataset of sequences were curated using FliC sequences from the tested *Salmonella* strains (Table S2). Alignments were performed in SeaView version 4.7 (Gouy *et al.*, 2010) using MUSCLE v3.8.31 (Edgar, 2004) using the default settings. Trees were calculated in SeaView, the model set as LG and 500 bootstrap samples using best of NNI and SPR for tree searching operations. Mapping and visualisation of datasets onto the tree was performed using itol (Letunic & Bork, 2011).

Protein sequence and domain analysis of each YSD1 ORF was performed using BLASTp (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and HMMer 2.31.0 Pfam (https://www.ebi.ac.uk/Tools/hmmer/) respectively. Structural predictions were performed using HHpred [(Zimmermann *et al.*, 2018); https://toolkit.tuebingen.mpg.de/#/tools/hhpred)]

and Phyre2.0 [(Kelley *et al.*, 2015);
http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index].

Protein purification and antibody production

The ORFs for YSD1_22, YSD1_25 and YSD1_29 were synthesized (genscript) with a C-terminal hexa-histidine tag, cloned into the protein expression vector pET21a (novagen) and transformed into *E. coli* C41. Cells were grown in 5 L of Terrific Broth (12 g tryptone, 24 g yeast extract, 4 mL glycerol, 2.31 g KH_2PO_4 and 12.84 g K_2HPO_4 per 1 L) at 37 °C until cultures reached an optical density (OD600) of 0.8 and protein expression was induced with 0.5 mM IPTG, shaking overnight at 18 °C. Cells were collected and lysed in lysis buffer (50 mM Tris pH 8.0, 400 mM NaCl, 2 mM $MgCl_2$ and 20 mM imidazole) using an Avestin cell press (3 passes). His-tagged proteins were first purified by Ni-affinity chromatography, with lysis buffer used for binding to the 5 mL nickel HisTrap HP column (GE Healthcare) and elution buffer (50 mM Tris pH 8.0, 400 mM NaCl and 1 M imidazole) to elute each protein respectively. Proteins were further purified by size exclusion chromatography using a HiLoad 16/600 Superdex 200 pg column (GE Healthcare) equilibrated in 25 mM Tris pH 8.0 and 200 mM NaCl. Polyclonal antibodies targeting YSD1_22, YSD1_25 and YSD1_29 were raised in mice (Wellcome Trust Sanger Institute, Cambridge, UK) by injecting each protein (200 ng) in phosphate buffered saline with Alum respectively.

SDS-PAGE and western blot analysis

YSD1 preparations were added to Pierce™ Lane Marker reducing sample buffer (ThermoFisher Scientific) and separated by a 4-12% Bis-Tris NuPAGE gel (ThermoFisher Scientific). Proteins from the purified YSD1 preparation (10^8 PFU) were subsequently stained with InstantBlue (Expedeon) for visualisation. YSD1-infected *S.* Typhi BRD948 lysates (10^8 PFU) were analysed by western immuno-blot to test the specificity of the antibodies raised against YSD1_22, YSD1_25 and YSD1_29 respectively. Antibodies were used at a dilution of 1:20,000 for YSD1_22 and 1:10,000 for YSD1_25 and YSD1_29.

Electron microscopy

For transmission electron microscopy (TEM) analysis of YSD1 and *S.* Typhi BRD948, 3 µL of sample was incubated on a freshly glow-discharged carbon-coated Formvar grid for 30 seconds. The sample was blotted from the grid and stained with 3 µL of 5% ammonium

molybdate for a further 30 seconds and blotted again. Grids were imaged using a 120 kV FEI Spirit BioTwin with a Tietz F4.15 CCD.

Immunogold-labelling of phage particles

All incubations and washes were performed on liquid drops on Parafilm at room temperature. Freshly glow-discharged carbon-coated Formvar grids were incubated on a drop of YSD1 stock solution for 3 seconds. Excess liquid was removed using Whatman filter paper and grids were immediately incubated on a drop of antibody solution (1:100 dilution in 5% foetal calf serum) for 20 minutes. Grids were rinsed briefly by transferring onto two consecutive drops of PBS with 0.01% Tween followed by a third drop of PBS only. Grids were then incubated on a drop containing 5 nm protein A-gold in 5% foetal calf serum for 10 minutes and subsequently washed as described above finishing on a drop of distilled water. Excess liquid was blotted away and grids were stained with 5% ammonium molybdate and 0.5% trehalose and imaged as described above.

Mass spectrometry

Samples were reduced with tris(2-carboxyethyl)phosphine, alkylated with iodoaceamide, and then digested with mass spectrometry grade trypsin (Pierce™, ThermoFisher Scientific) at 37°C for 18 h. After acidification with 5% formic acid, the sample was analysed by LC-MS/MS on an Ultimate 3000 RSLCnano System (Dionex) coupled to an LTQ Orbitrap Velos (ThermoFisher Scientific) mass spectrometer equipped with a nanospray source. The peptides were first loaded and desalted on a PepMap C18 trap (0.1 mm id x 20 mm, 5µm) then separated on a PepMap 75 µm id x 50 cm column (2 µm) over a 120 min linear gradient of 4 – 36% acetonitrile/0.1% formic acid. The LTQ Orbitrap Velos was operated in the "top 10" data-dependant acquisition mode, with the preview mode of FT master scan was disabled. The Orbitrap full scan was set at m/z 380 – 1500 with resolution set at 30,000 at m/z 400 and AGC at $1x10^6$ with a maximum injection time at 200 msec. The ten most abundant multiply-charged precursor ions, with a minimal signal above 3000 counts, were dynamically selected for CID fragmentation (MS/MS) in the LTQ ion trap, which has the AGC set at 5000 with the maximum injection time at 100 msec. The dynamic exclusion was set at ± 20 ppm for 60 sec.

Raw data was processed in Proteome Discoverer 1.4 (ThermoFisher Scientific) using the Sequest HT search engine against a combined protein database covering the host *Salmonella* Typhi (Uniprot; August 2016), YSD1 and the in-house contaminate database. The precursor

mass tolerance was set at 20 ppm, and fragment ions detected at 0.6 Da. Dynamic modifications were set as acetyl (protein N-term), carbamidomethyl (C), deamidated (NQ) and Oxidation (M). The False Discovery Rates (q-value) used were set at both strict (0.01) and relaxed (0.05) parameters for analysis. Only peptides at high confidence were selected for protein groups.

Small Angle X-ray Scattering

Small angle X-ray scattering (SAXS) was performed using Coflow SEC-SAXS at the Australian Synchrotron (Kirby *et al.*, 2016). Purified YSD1_29$^{373\text{-}1296}$ was analysed at a pre-injection concentration of 10 mg mL^{-1}. Scattering was collected over a q range of 0.0 to 0.3 Å^{-1}. A buffer blank for each SEC-SAXS run was prepared by averaging 10-20 frames pre or post protein elution. Scattering curves from peaks corresponding to YSD1_29$^{373\text{-}1296}$ were then buffer subtracted and scaled across the elution peak and compared for inter-particle effects. Identical curves (5-10) from elution were then averaged to provide curves for analysis. Data was analysed using the Primus package, Scatter and Dammif modeller (Konarev *et al.*, 2003).

Structural modelling of flagellar filament

As the structure of the *S.* Typhimurium flagellar filament (PDB: 3A5X) was deposited as a single subunit, the rest of the filament was modelled based on the helical symmetry of the *B. subtillis* flagellar filament [PDB:5WJT, (Wang *et al.*, 2017)] in Chimera 1.13.1 [(Pettersen *et al.*, 2004); http://www.rbvi.ucsf.edu/chimera/]. The 3A5X monomer was aligned to a subunit of the *B. subtillis* flagellar filament and the helical symmetry was applied to generate additional subunits. Residues in domains D2 (176-190, 285-402) and D3 (191-284) are highlighted in orange and green respectively. Structural models of different FliC variants were made using the one-to-one threading function of Phyre2.0 (Kelley *et al.*, 2015) using the *S.* Typhimurium FliC structure (PDB:3A5X) as a template. Electrostatics were calculated using the Coulumbic Surface Colouring using Chimera 1.13.1.

ACKNOWLEDGEMENTS

We acknowledge the Australian Synchrotron for access to beamlines MX1 and MX2 for X-ray crystallographic analysis (CAP11027) and SAXS-WAXS for X-ray solution scattering (M12480), and the the Sanger Institute DNA pipeline team for support in sequencing and assembling the genome of YSD1. We would also like to thank Alex Bateman of the EBI-EMBL for help with the use of the jackhammer software in detecting protein domains.

Research was supported by Program Grant 1092262 from the National Health and Medical Research Council of Australia (NHMRC), The Wellcome Trust and the Cambridge Biomedical Research Centre (National Institute for Medical Research, United Kingdom). R.G. is a Sir Henry Wellcome Fellow (106077/Z/14/Z), T.L. is an Australian Research Council Laureate Fellow (FL130100038).

AUTHOR CONTRIBUTIONS

RAD, DP, SD, DG, CC, JH, FL, RG, KH, JC, LY designed and carried out analysis. DP, DG, RG, JS, FS, SC, FC, RAS provided expertise to analyses. DG, RAS and TL supervised experimental work and evaluated data. RAD, DP, GD and TL evaluated results and wrote manuscript.

CONFLICT OF INTEREST

The authors declare no competing financial interests.

DATA AVAILABILITY

The data that support the findings of this study are openly available in Genbank at https://www.ncbi.nlm.nih.gov/genbank/, reference number LR026998.

REFERENCES

Ackermann, H.W. & D. Prangishvili, (2012) Prokaryote viruses studied by electron microscopy. *Arch Virol* **157**: 1843-1849.

Baldvinsson, S.B., M.C. Sorensen, C.S. Vegge, M.R. Clokie & L. Brondsted, (2014) Campylobacter jejuni motility is required for infection of the flagellotropic bacteriophage F341. *Appl Environ Microbiol* **80**: 7096-7106.

Berg, H.C. & R.A. Anderson, (1973) Bacteria swim by rotating their flagellar filaments. *Nature* **245**: 380-382.

Bonifield, H.R. & K.T. Hughes, (2003) Flagellar phase variation in Salmonella enterica is mediated by a posttranscriptional control mechanism. *J Bacteriol* **185**: 3567-3574.

Bradley, D.E., C.J. Douglas & J. Peschon, (1984) Flagella-specific bacteriophages of Agrobacterium tumefaciens: demonstration of virulence of nonmotile mutants. *Can J Microbiol* **30**: 676-681.

Casjens, S.R. & R.W. Hendrix, (2015) Bacteriophage lambda: Early pioneer and still relevant. *Virology* **479-480**: 310-330.

Charbit, A., (2003) Maltodextrin transport through lamb. *Front Biosci* **8**: s265-274.

Chatterjee, S. & E. Rothenberg, (2012) Interaction of bacteriophage l with its E. coli receptor, LamB. *Viruses* **4**: 3162-3178.

Edgar, R.C., (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792-1797.

Evans, L.D., S. Poulter, E.M. Terentjev, C. Hughes & G.M. Fraser, (2013) A chain mechanism for flagellum growth. *Nature* **504**: 287-290.

Finn, R.D., J. Clements & S.R. Eddy, (2011) HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res* **39**: W29-37.

Frankel, G., S.M. Newton, G.K. Schoolnik & B.A. Stocker, (1989) Intragenic recombination in a flagellin gene: characterization of the H1-j gene of Salmonella typhi. *EMBO J* **8**: 3149-3152.

Fukuda, A., K. Miyakawa, H. Iba & Y. Okada, (1976) A flagellotropic bacteriophage and flagella formation in Caulobacter. *Virology* **71**: 583-592.

Geiben-Lynn, R., K. Sauber & F. Lutz, (2001) Flagellin inhibits Myoviridae phage phiCTX infection of Pseudomonas aeruginosa strain GuA18: purification and mapping of binding site. *Arch Microbiol* **176**: 339-346.

Gonzalez, F., R.F. Helm, K.M. Broadway & B.E. Scharf, (2018) More than Rotating Flagella: Lipopolysaccharide as a Secondary Receptor for Flagellotropic Phage 7-7-1. *J Bacteriol* **200**.

Gouy, M., S. Guindon & O. Gascuel, (2010) SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27**: 221-224.

Grose, J.H. & S.R. Casjens, (2014) Understanding the enormous diversity of bacteriophages: the tailed phages that infect the bacterial family Enterobacteriaceae. *Virology* **468-470**: 421-443.

Guerrero-Ferreira, R.C., P.H. Viollier, B. Ely, J.S. Poindexter, M. Georgieva, G.J. Jensen & E.R. Wright, (2011) Alternative mechanism for bacteriophage adsorption to the motile bacterium Caulobacter crescentus. *Proc Natl Acad Sci U S A* **108**: 9963-9968.

Guibourdenche, M., P. Roggentin, M. Mikoleit, P.I. Fields, J. Bockemuhl, P.A. Grimont & F.X. Weill, (2010) Supplement 2003-2007 (No. 47) to the White-Kauffmann-Le Minor scheme. *Res Microbiol* **161**: 26-29.

Harvey, H., J. Bondy-Denomy, H. Marquis, K.M. Sztanko, A.R. Davidson & L.L. Burrows, (2018) Pseudomonas aeruginosa defends against phages through type IV pilus glycosylation. *Nat Microbiol* **3**: 47-52.

Hendrix, R.W., C.C. Ko, D. Jacobs-Sera, G.F. Hatfull, M. Erhardt, K.T. Hughes & S.R. Casjens, (2015) Genome Sequence of Salmonella Phage chi. *Genome Announc* **3**.

Jollick, J.D. & B.L. Wright, (1974) A flagella specific bacteriophage for caulobacter. *J Gen Virol* **22**: 197-205.

Katsamba, P. & E. Lauga, (2019) Hydrodynamics of bacteriophage migration along bacterial flagella. *Physical Review Fluids* **4**: 013101.

Kazaks, A., A. Dislers, G. Lipowsky, V. Nikolajeva & K. Tars, (2012) Complete genome sequence of the Enterobacter cancerogenus bacteriophage Enc34. *J Virol* **86**: 11403-11404.

Kelley, L.A., S. Mezulis, C.M. Yates, M.N. Wass & M.J. Sternberg, (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**: 845-858.

Kirby, N., N. Cowieson, A.M. Hawley, S.T. Mudie, D.J. McGillivray, M. Kusel, V. Samardzic-Boban & T.M. Ryan, (2016) Improved radiation dose efficiency in solution SAXS using a sheath flow sample environment. *Acta Crystallogr D Struct Biol* **72**: 1254-1266.

Konarev, P.V., V.V. Volkov, A.V. Sokolova, M.H.J. Koch & D.I. Svergun, (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography* **36**: 1277-1282.

Kropinski, A.M., A. Van den Bossche, R. Lavigne, J.P. Noben, P. Babinger & R. Schmitt, (2012) Genome and proteome analysis of 7-7-1, a flagellotropic phage infecting Agrobacterium sp H13-3. *Virol J* **9**: 102.

Lederberg, E.M., (1951) Lysogenicity in E-coli K-12. In: Genetics. 428 EAST PRESTON ST, BALTIMORE, MD 21202, pp. 560-560.

Letunic, I. & P. Bork, (2011) Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res* **39**: W475-478.

Lovett, P.S., (1972) PBPI: a flagella specific bacteriophage mediating transduction in Bacillus pumilus. *Virology* **47**: 743-752.

Macnab, R.M., (1977) Bacterial flagella rotating in bundles: a study in helical geometry. *Proc Natl Acad Sci U S A* **74**: 221-225.

McCutcheon, J.G., D.L. Peters & J.J. Dennis, (2018) Identification and Characterization of Type IV Pili as the Cellular Receptor of Broad Host Range Stenotrophomonas maltophilia Bacteriophages DLP1 and DLP2. *Viruses* **10**.

Merino, S., S. Camprubi & J.M. Tomas, (1990) Isolation and characterization of bacteriophage PM3 from Aeromonas hydrophila the bacterial receptor for which is the monopolar flagellum. *FEMS Microbiol Lett* **57**: 277-282.

Meynell, E.W., (1961) A phage, phi chi, which attacks motile bacteria. *J Gen Microbiol* **25**: 253-290.

Minamino, T., (2014) Protein export through the bacterial flagellar type III export pathway. *Biochim Biophys Acta* **1843**: 1642-1648.

Moreno Switt, A.I., R.H. Orsi, H.C. den Bakker, K. Vongkamjan, C. Altier & M. Wiedmann, (2013) Genomic characterization provides new insight into Salmonella phage diversity. *BMC Genomics* **14**: 481.

Nobrega, F.L., M. Vlot, P.A. de Jonge, L.L. Dreesens, H.J.E. Beaumont, R. Lavigne, B.E. Dutilh & S.J.J. Brouns, (2018) Targeting mechanisms of tailed bacteriophages. *Nat Rev Microbiol* **16**: 760-773.

Onmus-Leone, F., J. Hang, R.J. Clifford, Y. Yang, M.C. Riley, R.A. Kuschner, P.E. Waterman & E.P. Lesho, (2013) Enhanced de novo assembly of high throughput pyrosequencing data using whole genome mapping. *PLoS One* **8**: e61762.

Pate, J.L., S.J. Petzold & T.H. Umbreit, (1979) Two flagellotropic phages and one pilus-specific phage active against Asticcacaulis biprosthecum. *Virology* **94**: 24-37.

Pettersen, E.F., T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng & T.E. Ferrin, (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**: 1605-1612.

Phothaworn, P., M. Dunne, R. Supokaivanich, C. Ong, J. Lim, R. Taharnklaew, M. Vesaratchavest, R. Khumthong, O. Pringsulaka, P. Ajawatanawong, J. Klumpp, N. Brown, M. Imam, M.R.J. Clokie, E.E. Galyov & S. Korbsrisate, (2019) Characterization of Flagellotropic, Chi-Like Salmonella Phages Isolated from Thai Poultry Farms. *Viruses* **11**.

Pickard, D., N.R. Thomson, S. Baker, J. Wain, M. Pardo, D. Goulding, N. Hamlin, J. Choudhary, J. Threfall & G. Dougan, (2008) Molecular characterization of the Salmonella enterica serovar Typhi Vi-typing bacteriophage E1. *J Bacteriol* **190**: 2580-2587.

Pickard, D., A.L. Toribio, N.K. Petty, A. van Tonder, L. Yu, D. Goulding, B. Barrell, R. Rance, D. Harris, M. Wetter, J. Wain, J. Choudhary, N. Thomson & G. Dougan, (2010) A conserved acetyl esterase domain targets diverse bacteriophages to the Vi capsular receptor of Salmonella enterica serovar Typhi. *J Bacteriol* **192**: 5746-5754.

Potter, S.C., A. Luciani, S.R. Eddy, Y. Park, R. Lopez & R.D. Finn, (2018) HMMER web server: 2018 update. *Nucleic Acids Res* **46**: W200-W204.

Ravid, S. & M. Eisenbach, (1983) Correlation between bacteriophage chi adsorption and mode of flagellar rotation of Escherichia coli chemotaxis mutants. *J Bacteriol* **154**: 604-611.

Rossmann, F.M. & M. Beeby, (2018) Insights into the evolution of bacterial flagellar motors from high-throughput in situ electron cryotomography and subtomogram averaging. *Acta Crystallogr D Struct Biol* **74**: 585-594.

Samuel, A.D., T.P. Pitta, W.S. Ryu, P.N. Danese, E.C. Leung & H.C. Berg, (1999) Flagellar determinants of bacterial sensitivity to chi-phage. *Proc Natl Acad Sci U S A* **96**: 9863-9866.

Sanger, F., A.R. Coulson, G.F. Hong, D.F. Hill & G.B. Petersen, (1982) Nucleotide sequence of bacteriophage lambda DNA. *J Mol Biol* **162**: 729-773.

Schade, S.Z., J. Adler & H. Ris, (1967) How bacteriophage chi attacks motile bacteria. *J Virol* **1**: 599-609.

Schreiber, F., S. Kay, G. Frankel, S. Clare, D. Goulding, E. van de Vosse, J.T. van Dissel, R. Strugnell, G. Thwaites, R.A. Kingsley, G. Dougan & S. Baker, (2015) The Hd, Hj, and Hz66 flagella variants of Salmonella enterica serovar Typhi modify host responses and cellular interactions. *Sci Rep* **5**: 7947.

Sertic, V. & N. Boulgakov, (1936) Bactériophages spécifiques pour des variétés bactériennes flagellées. *Compt. Rend. Soc. Biol* **123**: 887-888.

Shea, T.B. & E. Seaman, (1984) SP3: a flagellotropic bacteriophage of Bacillus subtilis. *J Gen Virol* **65 (Pt 11)**: 2073-2076.

Shin, H., J.H. Lee, H. Kim, Y. Choi, S. Heu & S. Ryu, (2012) Receptor diversity and host interaction of bacteriophages infecting Salmonella enterica serovar Typhimurium. *PLoS One* **7**: e43392.

Tacket, C.O., M.B. Sztein, G.A. Losonsky, S.S. Wasserman, J.P. Nataro, R. Edelman, D. Pickard, G. Dougan, S.N. Chatfield & M.M. Levine, (1997) Safety of live oral Salmonella typhi vaccine strains with deletions in htrA and aroC aroD and immune response in humans. *Infect Immun* **65**: 452-456.

Vieira, G., H. de Lencastre & L. Archer, (1989) Restriction analysis of PBS 1-related phages. *Arch Virol* **106**: 121-126.

Wang, F., A.M. Burrage, S. Postel, R.E. Clark, A. Orlova, E.J. Sundberg, D.B. Kearns & E.H. Egelman, (2017) A structural model of flagellar filament switching across multiple bacterial species. *Nat Commun* **8**: 960.

Yonekura, K., S. Maki-Yonekura & K. Namba, (2003) Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature* **424**: 643-650.

Zhilenkov, E.L., V.M. Popova, D.V. Popov, L.Y. Zavalsky, E.A. Svetoch, N.J. Stern & B.S. Seal, (2006) The ability of flagellum-specific Proteus vulgaris bacteriophage PV22 to interact with Campylobacter jejuni flagella in culture. *Virol J* **3**: 50.

Zimmermann, L., A. Stephens, S.Z. Nam, D. Rau, J. Kubler, M. Lozajic, F. Gabler, J. Soding, A.N. Lupas & V. Alva, (2018) A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *J Mol Biol* **430**: 2237-2243.

FIGURE LEGENDS

Figure 1. Morphological assessment of phage YSD1. (A) A sample of YSD1 phage was purified from infected *S.* Typhi BRD948 and analysed by transmission electron microscopy (TEM). The scale bar is set at 100 nm. Zooming in on the tail (inset) revealed the long flexible tail-fibre characteristic of Chi-like phage. (B) YSD1 architecture, drawn to scale. (C) YSD1-infected *S.* Typhi BRD948 were analysed by TEM. The scale bar is set at 100 nm. In several views, long tail fibres from YSD1 were evident interacting with the bacterial flagellum (inset). (D) TEM sections of YSD1-infected *S.* Typhi BRD948 that show phage at the base of a flagellum. The scale bar is set at 100 nm or 200 nm, as indicated. (E) Phylogeny depicting the sequence diversity in flagellins across the family of *Enterobacteriaceae*. The inner circle and branches are color-coded to designate grouped bacterial genera, as indicated in the key. The scale bar represents the number of amino acid substitutions per site. The position of the selected *Salmonella* FliC proteins are further highlighted (black bars) to indicate those species/serovars that were determined to be sensitive to YSD1.

Figure 2. Summary of the genome of Chi-like phage YSD1. (A) The YSD1 genome was ordered based upon the cohesive ends identified in phage Chi, yielding a pair-wise sequence identity of ~ 97%. Conserved gene homologs are shown in green, distinct genes are shown in yellow. The genes are numbered, such that the gene encoding the major capsid protein YSD1_17 corresponds to the gene encoding the major capsid protein gp17 in Chi phage. (B) BLAST analysis, structural predictions and genome synteny were used to summarize the putative function-based regions of the YSD1 genome.

Figure 3. Architectural arrangement of major proteins that form the tail-tube and fibre of YSD1. (A) Purified YSD1 virions (~1.0 x 10^8 PFU) were analysed by SDS-PAGE and the gel stained with Coomassie blue. The migration position of molecular weight markers is shown. (B) YSD1-infected *S.* Typhi BRD948 extracts were prepared and separated by SDS-PAGE. Immunoblot analysis with antibodies raised against recombinant forms of the indicated proteins. (C) Samples of YSD1 were subject to immunogold labelling using antibodies to YSD1_22. The multiple micrographs are representative of the localizations observed. (D) Samples of YSD1 were subject to immunogold labelling using antibodies to YSD1_25. The multiple micrographs are representative of the localizations observed. (E) Samples of YSD1 were subject to immunogold labelling using antibodies to YSD1_29. The multiple micrographs are representative of the localizations observed. (F) The graphical output from predictions of native disorder using several tools, with disorder probability plotted against the protein sequence of YSD1_29. (G) YSD1_29 was purified as a monomeric species and analysed by small-angle X-ray scattering (SAXS).

Figure 4. Structural features in the *Salmonella* flagellum. (A) Cartoon of flagellum [modified from (Evans *et al.*, 2013, Rossmann & Beeby, 2018)]. The various protein components are indicated. Either FliC or FljB can form the main portion of the filament. (B) A stock of YSD1 (10^{10} PFU ml^{-1}) was serially diluted from 10^{-1} to 10^{-6} and spotted onto a lawn of either wild-type *S.* Typhimurium, or the *ΔfliC*, *ΔfljB* mutant of *S.* Typhimurium. (C) Phylogenetic analysis of flagellin sequences from these *Salmonella* serovars. As indicated, this grouping also segregates the serovars according to their *H* antigen designations (antigenic formulae). The colours indicate common antigenic formulae of the phase 1 Flagellin protein. The scale bar represents the number of amino acid substitutions per site. (D) The FliC monomer has four domains organized in a bent hair-pin structure, D0 (red), D1 (blue), D2 (yellow) and D3 (green). (E) Sequence conservation analysis across the serovars, mapped

onto this structure. The D2 and D3 domains show the most sequence variation (blue) between serovars. (F) A structural model of a FliC filament from *S.* Typhimurium, based on a cryo-EM structure [5WJT; (Wang *et al.*, 2017)]. The model shows the central protein export channel, and the radial disposition of the flagellin D2 (orange) and D3 (green) domains in the context of the filament. (G) A cross section of the structural model of the FliC filament showing the ability of a strand of DNA to fit within the central cavity of the flagellum.

TABLE 1. Annotation of YSD1 open-reading frames.

Gene Name	BLAST/HMMer Pfam Description	Phyre 2.0/HHpred prediction
YSD1_1	helix-turn-helix XRE family of transcriptional regulators	Putative XRE family of transcriptional regulator, lambda repressor-like DNA-binding domain
YSD1_2	DNA primase	DNA primase
YSD1_3	Putative transcriptional regulator	Lambda repressor-like DNA-binding domain, GalR/LacI-like-bacterial regulator
YSD1_4	Hypothetical protein, RecT-like protein	-
YSD1_5	Hypothetical protein, Cas4-like protein	-
YSD1_6	Hypothetical protein, Cas4-like protein, DUF2800	Crispr-associated exonuclease, ATP-dependent helicase/DNase subunit B
YSD1_7	Hypothetical protein, DUF2815, ssDNA-binding protein	ssDNA-binding protein
YSD1_8	DNA polymerase I	DNA polymerase I
YSD1_9	VRR-NUC domain containing protein	VRR_NUC domain
YSD1_10	DNA helicase, SNF2 domain-containing protein	Recombination/DNA-binding protein
YSD1_11	Putative Terminase small-subunit TerS	Terminase gpNU1 subunit domain
YSD1_12	Putative Terminase large-subunit TerL	Phage Terminase large subunit
YSD1_13	Putative head-to-tail joining protein W, gpW superfamily	Head-to-tail joining protein W, gpW
YSD1_14	Phage portal protein	Portal protein, head-to-tail joining protein
YSD1_15	Capsid maturation protease ClpP	Protease 4, ATP-dependent Clp protease proteolytic subunit
YSD1_16	Decorative head protein D/Auxiliary protein	Head decoration protein D
YSD1_17	Major capsid protein E	Putative capsid protein
YSD1_18	Hypothetical protein	-
YSD1_19	Hypothetical protein	Lambda gpFII like protein, head-tail joining protein
YSD1_20	Hypothetical protein	-

YSD1_21	Hypothetical protein	Minor tail protein U
YSD1_22	Major tail protein, bacterial Ig-like domain-containing protein	Tail tube protein
YSD1_23	Tail assembly chaperone	-
YSD1_24	Tape measure protein	-
YSD1_25	Putative distal tail protein	Distal tail protein
YSD1_26	Putative tail assembly protein, DUF2163	Putative prophage mu tail protein
YSD1_27	Hypothetical protein, putative tail assembly protein	-
YSD1_28	Hypothetical protein, putative tail assembly protein	-
YSD1_29	Tail fibre protein, minor tail protein, Phage-tail_3 superfamily	Baseplate component of bacteriophage mu
YSD1_30	DUF2793	-
YSD1_31	Hypothetical protein	CIP1 like protein, Glucuronan lyase, Alginate lyase, Sialidase
YSD1_32	Hypothetical protein	Lectins/glucanases, CIP1 like protein, Glucuronan lyase, Alginate lyase, Sialidase
YSD1_33	Hypothetical protein	Glucuronan lyase a, cellulose induced protein, cip1, Concanavalin A-like lectins/glucanases
YSD1_34	Hypothetical protein	CIP1 like protein, Glucuronan lyase, Alginate lyase, Sialidase
YSD1_35	Hypothetical protein	-
YSD1_36	Putative endolysin 2, lysis protein B, holin	-
YSD1_37	Hypothetical protein, putative endolysin 1, lysis protein A	Putative peptidoglycan hydrolase
YSD1_38	Hypothetical protein, endolysin like protein	-
YSD1_39	Hypothetical protein, TIR_2 superfamily, nucleoside 2-deoxyribosyltransferase	N-(deoxy)ribosyltransferase
YSD1_40	Hypothetical protein (found in SPN19, iEPS5 and Enc34 phages)	Transcriptional regulatory protein, DNA polymerase III alpha subunit
YSD1_41	Hypothetical protein	-
YSD1_42	Hypothetical protein	-

YSD1_43	Hypothetical protein, putative DNA-directed DNA polymerase protein	-
YSD1_44	Hypothetical protein	-
YSD1_45	Hypothetical protein	-
YSD1_46	Hypothetical protein, FMT_C_like superfamily	-
YSD1_47	Hypothetical protein	-
YSD1_48	Hypothetical protein	-
YSD1_49	Putative N-6-adenine-methyltransferase, DAM superfamily	S-adenosyl-L-methionine-dependent methyltransferases
YSD1_50	Hypothetical protein	-
YSD1_51	Hypothetical protein	-
YSD1_52	Hypothetical protein	-
YSD1_53	Hypothetical protein, putative endolysin	-
YSD1_54	Hypothetical protein, putative tail fibre protein	-
YSD1_55	Exonuclease, putative recombination-associated protein RdgC	Recombination-associated protein RdgC
YSD1_56	Hypothetical protein	-
YSD1_57	Hypothetical protein	DNA binding protein
YSD1_58	Hypothetical protein	-
YSD1_59	Hypothetical protein	-
YSD1_60	Hypothetical protein	-
YSD1_61	Hypothetical protein	-
YSD1_62	Hypothetical protein	-
YSD1_63	Hypothetical protein	-
YSD1_64	Hypothetical protein	-
YSD1_65	Hypothetical protein, resolvase, holin	-
YSD1_66	Hypothetical protein, putative DNA binding protein	-

YSD1_67	Hypothetical protein	-
YSD1_68	Hypothetical protein	-
YSD1_69	Hypothetical protein, Transglut_core superfamily	-
YSD1_70	Terminase large subunit, Hypothetical protein	-
YSD1_71	Hypothetical protein, putative outer capsid protein	-

Footnote: All proteins characterised as hypothetical have homologs present in other Chi-like phages.

TABLE 2. Mass spectrometry analysis of YSD1 virions.

Protein	Description	MW (kDa)	% Identity to homolog in Chi
YSD1_24	Tape measure protein	153.7	99%
YSD1_29	Minor tail protein	143.1	99%
YSD1_35	Hypothetical protein	78.4	96%
YSD1_25	Putative distal tail protein	62.9	99%
YSD1_14	Portal protein	62.2	99%
YSD1_15	Capsid maturation protein	46.3	99%
YSD1_34	Hypothetical protein, putative lyase	44.2	99%
YSD1_67	Hypothetical protein, putative DNA binding protein	43.2	94%
YSD1_22	Major tail protein	40.3	99%
YSD1_17	Major capsid protein	39.9	99%
YSD1_55	Putative Exonuclease	39.8	99%
YSD1_33	Hypothetical protein, putative lyase	36.9	100%
YSD1_31	Hypothetical protein, putative lyase	36.4	100%
YSD1_32	Hypothetical protein, putative lyase	34.7	100%
YSD1_48	Hypothetical protein	30.8	97%
YSD1_52	Hypothetical protein	27.9	99%
YSD1_49	Putative N-6-adenine-methyltransferase	26.1	92%
YSD1_20	Hypothetical protein	22.5	100%
YSD1_7	Hypothetical protein, DUF2815, Single-stranded DNA binding protein	22.5	98%
YSD1_21	Hypothetical protein	18.8	100%
YSD1_16	Putative decorative head protein	14.4	100%
YSD1_4	Hypothetical protein, RecT-like protein	14.1	99%
YSD1_59	Hypothetical protein	13.4	100%
YSD1_18	Hypothetical protein	10.6	100%

TABLE 3. Neutralization of YSD1 infectivity with antibodies.

Condition	Antibody raised against:			
	YSD1_22	**YSD1_25**	**YSD1_29**	**YSD1_16**
Phage alone	10^{-6}	10^{-6}	10^{-6}	10^{-6}
Antibody alone	-	-	-	-
Antibody 1:10 + phage	10^{-4}	10^{-5}	-	10^{-5}
Antibody 1:100 + phage	10^{-4}	10^{-5}	10^{-4}	10^{-6}
Antibody 1:1000 + phage	10^{-5}	10^{-5}	10^{-5}	10^{-6}
Antibody 1:10000 + phage	10^{-5}	10^{-5}	10^{-5}	10^{-6}

Footnotes:

YSD1 was pre-treated with the indicated antibodies.

The lowest concentration of phage stock which generated visible plaques/zones of clearance is indicated.

TABLE 4. Efficiency of plating (EOP) data.

Salmonella Strain	Salmonella Subspecies	Phase 1 Antigen (h)	Phase 2 Antigen (h)	'R' phase	YSD1 infection	EOP
S. Miami	Group I	a	1,5	-	+	10^{-3}
S. Abony	Group I	b	e, n, x	-	+	10^{-2}
S. Choleraesuis	Group I	c	1,5	-	+	10^{-2}
S. Abortus ovis	Group I	c	1,6	-	+	10^{-1}
S. Typhi BRD948 Hd	Group I	d	-	-	+	1
S. Stanley	Group I	d	1,2	-	+	10^{-1}
S. Schwarzengrund	Group I	d	1,7	-	+	10^{-1}
S. Duisburg	Group I	d	e, n, z_{15}	e, h	+	10^{-4}
S. Typhi BRD948 Hj	Group I	-	-	j	+	1
S. Abortus equi	Group I	-	e, n, x	-	+	1
S. Anatum	Group I	e, h	1,6	z_{64}	-	-
S. Newport	Group I	e, h	1,2	z_{67}, z_{78}	-	-
S. Reading	Group I	e, h	1,5	R1	-	-
S. Saintpaul	Group I	e, h	1,2	-	-	-
S. Agona	Group I	f, g, s	1,2	z_{27}, z_{45}	-	-

S. Dublin	Group I	g, p	-	-	-	-
S. Enteritidis	Group I	g, m	-	-	-	-
S. Montevideo	Group I	g, m, s	-	-	-	-
S. Kentucky	Group I	i	z_6	-	+	10^{-4}
S. Typhimurium	Group I	i	1,2	-	+	10^{-1}
S. Heidelberg	Group I	r	1,2	-	+	10^{-1}
S. Infantis	Group I	r	1,5	R1, z_{37}, z_{45}, z_{49}	+	10^{-2}
S. Hadar	Group I	z_{10}	e, n, x	-	+	1
S. Javiana	Group I	1, z_{28}	1,5	R1	+	10^{-1}
S. Albany	Group I	z_4, z_{24}	-	z_{45}	+	10^{-2}
S. Typhi BRD948 Hz66	Group I	-	-	z_{66}	-	-
S. Gallinarum biovar Pullorum	Group I	No Flagella			-	-
S. Arizoniae	Group IIIa	z_4, z_{23}, z_{26}	-	Not specified	+	10^{-3}
S. Diarizonae	Group IIIb	e,n, z_{24}	-	Not specified	+	10^{-2}
S. bongori	Group V	Not specified	Not specified	Not specified	+	10^{-1}

SUPPLEMENTAL INFORMATION

Supporting Table S1: SAXS parameters for YSD1_29$^{373-1296}$

Supporting Table S2: List of *Salmonella* serovars and their corresponding FliC protein sequences used to screen YSD1 sensitivity

Supporting Table S3: List of Flagellin proteins from members of the *Enterobacteriaceae* family used for phylogenetic tree analysis

Supporting Figure S1: YSD1 infection of *S.* Typhi BRD948.
Infection assays in which different dilutions of YSD1 were spotted onto a lawn of (A) *S.* Typhi BRD948 and *S.* Typhi BRD948 *ΔfliC* and (B) *S.* Typhi BRD948 and *S.* Typhi strain BA256 (*ΔtviB*).

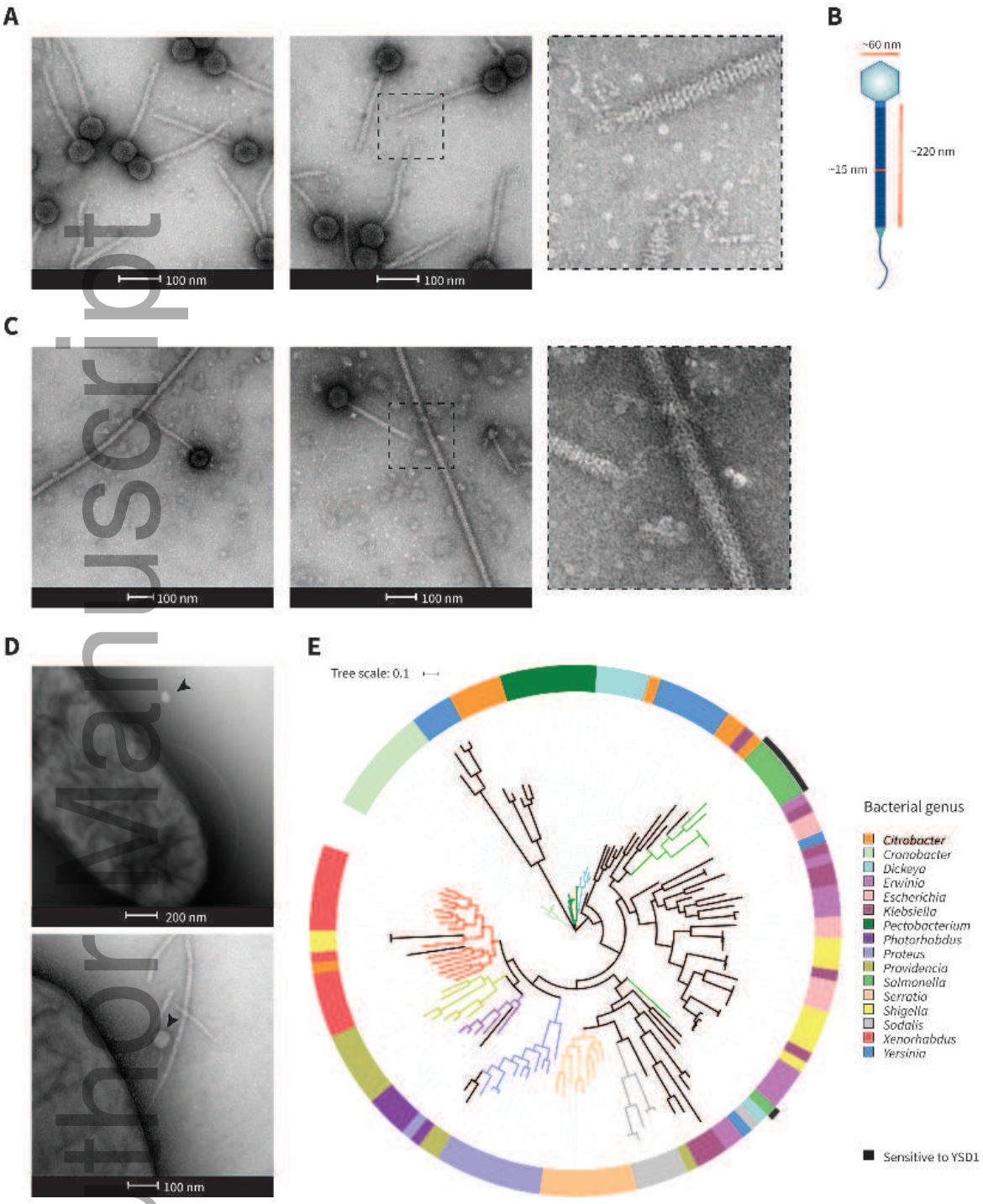
Supporting Figure S2: Distribution of gold particles after labelling of YSD1 tail structures.
To distinguish between the labelling of YSD1_29 and YSD1_25, the distance of each gold particle was measured from an equivalent central point on the YSD1 baseplate. The distance (in nm) from a total of 38 particles for each antibody were measured and represented by a box plot and unpaired t-test analysis ($p = 0.002$).

Supporting Figure S3: SAXS scattering analysis for YSD1_29$^{373\text{-}1296}$.
(A) YSD1_29 post size exclusion chromatography showing the YSD1_29$^{373\text{-}1296}$ protein (~105 kDa) analysed by SAXS. (B) X-ray scattering profile of YSD1_29$^{373\text{-}1296}$, (C) Kratky plot of YSD1_29$^{373\text{-}1296}$ scattering, showing a single domain protein with some flexibility, (D) P(r) plot of YSD1_29$^{373\text{-}1296}$ scattering showing that the protein adopts a highly elongated confirmation with maximum dimensions of 257 Å. (E) Poyrod-Deybe plot of YSD1_29$^{373\text{-}1296}$, (F) The fit of the YSD1_29$^{373\text{-}1296}$ (red line) model calculated with DAMMIF, shown in Fig. 3G, with the X-ray scattering of YSD1_29$^{373\text{-}1296}$ (black points).

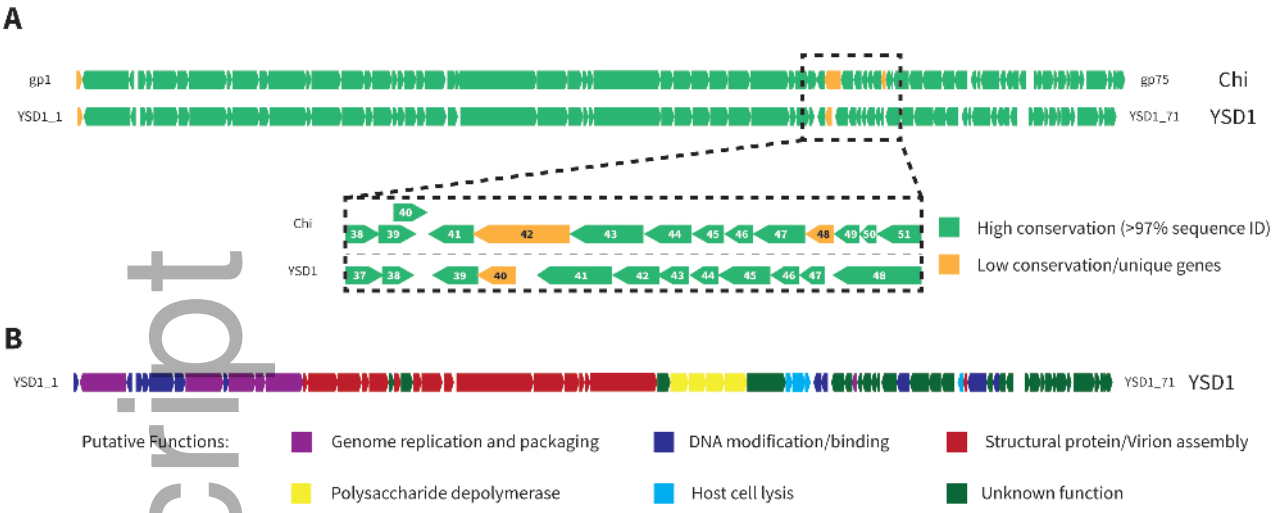
Supporting Figure S4: Inhibition of YSD1 infection by antibodies targeting the flexible tail fibre protein YSD1_29.
Spot test assay showing plaques produced by YSD1 pre-treated with antibodies targeting YSD1_29 (Table 3). Serial dilutions of YSD1 phage (10^{-1} to 10^{-6}) and the YSD1_29 antibody (10^{-1} to 10^{-4}) were prepared, mixed in a 1:1 ratio and incubated at 37 °C for 1 hour. The phage-antibody mix was then spotted onto a lawn of *S.* Typhimurium SL1344 and incubated overnight at 37 °C.

Supporting Figure S5: Sequence and structural modelling of FliC variants from selected *Salmonella* serovars.

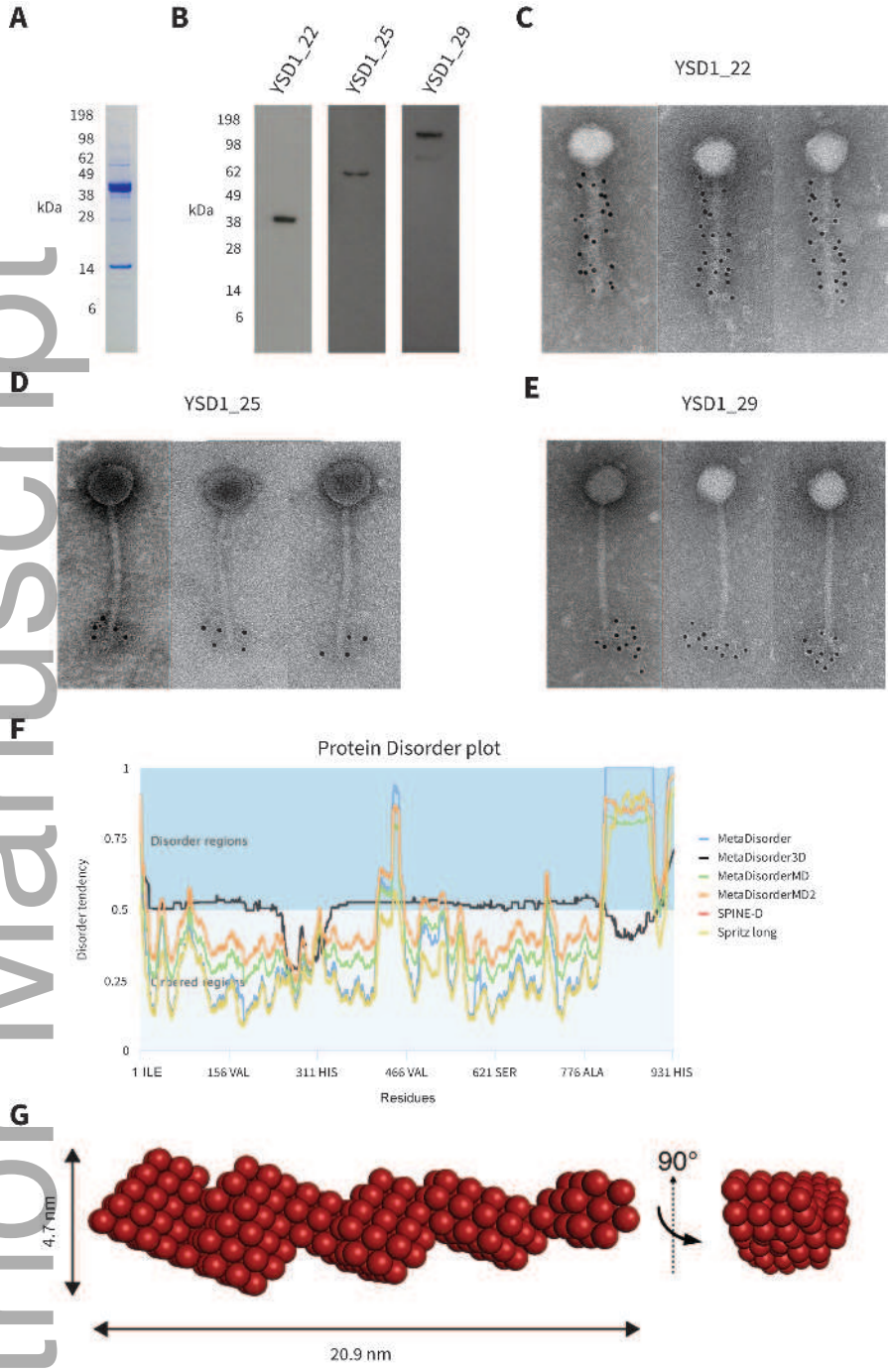
(A) Model of the FliC monomer from *S.* Typhi H*d* showing the residues (coloured in black) absent in the *S.* Typhi H*j* R-phase of FliC and alignment of *S.* Typhi H*d* and H*j* FliC protein sequences. (B) Surface mapping detailing the electrostatic properties of different modelled FliC proteins.



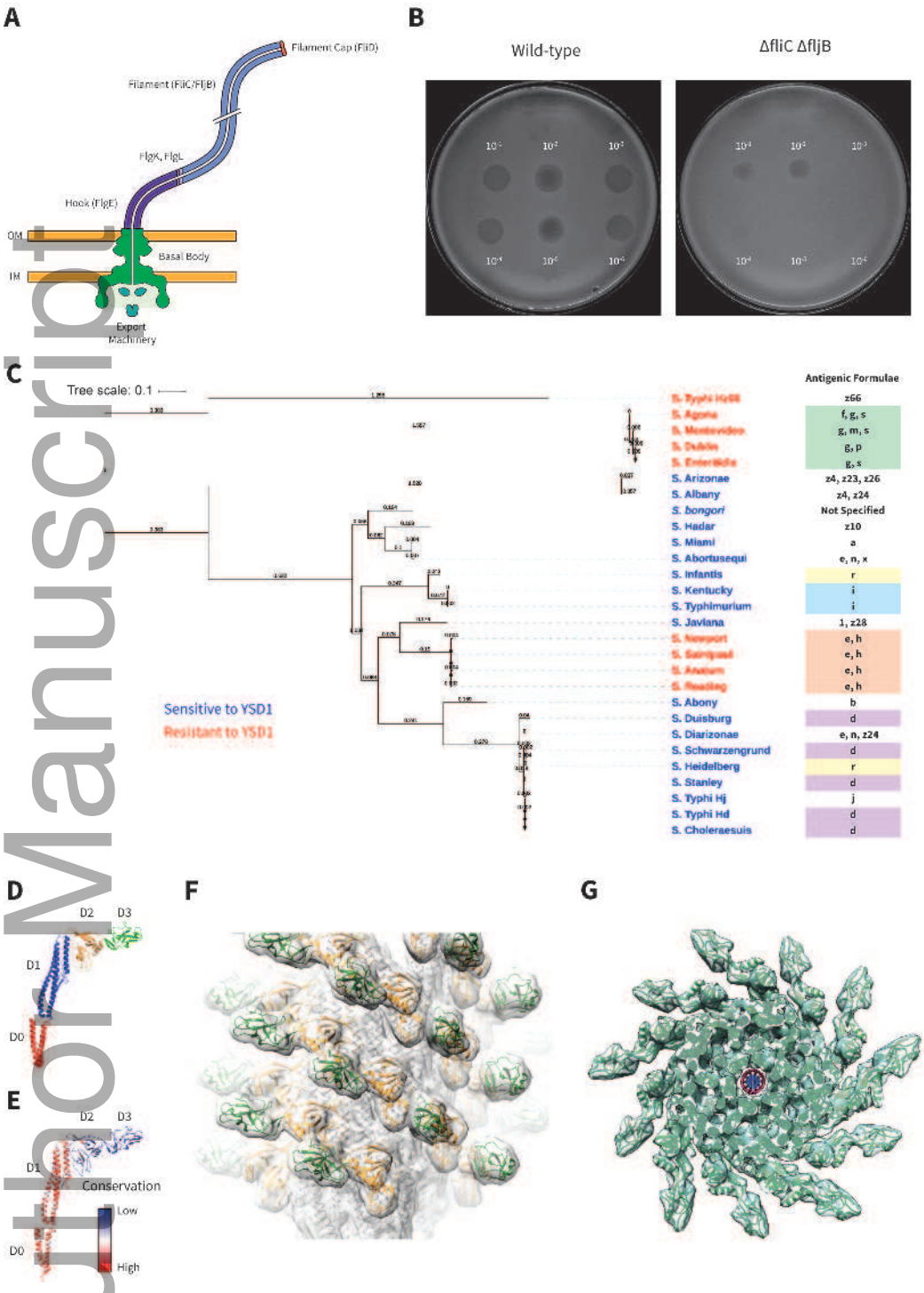
mmi_14396_f1.tif



mmi_14396_f2.tif



mmi_14396_f3.tif



mmi_14396_f4.tif