Characterization of the relationship between integrase, excisionase and antirepressor activities associated with a superinfecting Shiga toxin encoding bacteriophage

P. C. M. Fogg1, D. J. Rigden2, J. R. Saunders1, A. J. McCarthy1 and H. E. Allison1,*

1Microbiology Research Group, School of Biological Sciences, University of Liverpool and 2Structural Biology Group, School of Biological Sciences, University of Liverpool, BioSciences Building, Crown Street, Liverpool L69 7ZB, UK

Received August 24, 2010; Revised September 23, 2010; Accepted September 24, 2010

ABSTRACT

Shigatoxigenic Escherichia coli emerged as new food borne pathogens in the early 1980s, primarily driven by the dispersal of Shiga toxin-encoding lambdoid bacteriophages. At least some of these Stx phages display superinfection phenotypes, which differ significantly from lambda phage itself, driving through in situ recombination further phage evolution, increasing host range and potentially increasing the host’s pathogenic profile. Here, increasing levels of Stx phage Φ24B integrase expression in multiple lysogen cultures are demonstrated along with apparently negligible repression of integrase expression by the cognate CI repressor. The Φ24B int transcription start site and promoter region were identified and found to differ from in silico predictions. The unidirectional activity of this integrase was determined in an in situ, inducible tri-partite reaction. This indicated that Φ24B must encode a novel directionality factor that is controlling excision events during prophage induction. This excisionase was subsequently identified and characterized through complementation experiments. In addition, the previous proposal that a putative antirepressor was responsible for the lack of immunity to superinfection through inactivation of CI has been revisited and a new hypothesis involving the role of this protein in promoting efficient induction of the Φ24B prophage is proposed.

INTRODUCTION

Shigatoxigenic Escherichia coli (STEC) emerged in the early 1980s as the causal agents of a variety of clinical symptoms and sequelae ranging from mild diarrhea through to life threatening conditions such as hemolytic uremic syndrome (HUS) and thrombotic thrombocytopenic purpura (TTP) (1). The most prevalent STEC strains in the UK, North America and Japan belong to the infamous O157:H7 serogroup (2–4), though several other serogroups dominate mainland Europe and South America (5,6). The most notable non-O157 serogroups known to be strongly associated with human outbreaks are O26 (7), O111 (8) and O145 (9). STEC are characterized by carriage of numerous pathogenicity determinants, but it is the presence of Stx-phages, encoding the eponymous Shiga toxins (Stx) (10,11), which is without doubt the primary factor responsible for the most serious disease manifestations. Significantly, these Stx-phages also have the ability to convert non-pathogenic, commensal bacteria into toxin producers, thus augmenting disease progression (10,12,13).

Stx-phages are classified as lambdoid and, as such, are characteristically temperate and encode an immunity system analogous to bacteriophage lambda (14,15). Immediately after bacteriophage entry into a host cell, Stx-phage can follow either a lytic or lysogenic replication strategy, with the latter resulting in site-directed integration of the phage genome into the bacterial chromosome to generate a prophage, which later can be induced to enter the lytic replication route if exposed to an appropriate induction signal (16). Establishment of lysogeny requires bacteriophage expression of integrase (Int). Usually, the integrases encoded by lambdoid phages are

*To whom correspondence should be addressed. Tel: +44 151 795 4571; Fax: +44 151 795 4410; Email: hallison@liv.ac.uk
Present address:
P.C.M. Fogg, Department of Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia, Canada.

© The Author(s) 2010. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.5), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
tyrosine recombinases, which, during phage integration, form an Int/DNA intasome complex to catalyze unidirectional creation and resolution of a four way Holliday junction, ultimately yielding an integrated prophage (17). Integrase is the only phage-encoded protein required to facilitate this reaction. Once integrated, the lambda prophage produces a repressor (CI), which self-regulates its own promoter and inhibits all other phage gene expression in the lysogen plus any additional incoming homologous phage, thus maintaining stable lysogeny and precluding superinfection (15). However, contrary to expectations, various different Stx-phages have been described that are able to circumvent superinfection immunity at high frequency (18, 19). The Stx-phage, Φ24B, can integrate into a single host at least three times, consecutively, with increasing efficiency during each successive infection (20). Multiple infections of this nature may lead to amplified toxin production resulting from a greater stx gene copy number and increased opportunity for intracellular recombination with consequent evolutionary implications (21, 22). Φ24B also encodes a putative antirepressor gene, ant, which shares 82% C-terminal identity to the archetypal antirepressor of the Salmonella bacteriophage P22 (20). Under normal conditions, the P22 antirepressor is not expressed and superinfection events are blocked, but under conditions that enable its expression, the P22 antirepressor binds and inactivates the P22 repressor protein, CII (23–25). Furthermore, the P22 Ant is also effective against the homologous λ repressor CI (24). This Ant-mediated repressor inactivation is a potential mechanism to explain the lack of superinfection immunity in Φ24B (20).

Under certain conditions, such as DNA damage, an integrated phage may be induced to excise itself from the host genome and enter the lytic replication cycle. Activated RecA (part of the host SOS response) induces autodigestion of the repressor (26), allowing infective phage particles to be produced. Similarly to integration, bacteriophage excision is also catalyzed by integrase, though a phage-encoded excisionase is required additionally to complete the reaction (Xis or recombination directionality factor, RDF) (27, 28). RDFs may control the fate of the integrase-catalyzed reaction by directly interacting with the intasome, facilitating an allosteric change which favors excision (29), or alternatively, they may act as transcriptional activators, inducing expression of their cognate integrases thereby promoting phage excision (28). In most cases, an RDF/Xis factor is necessary to enable excision of an integrated stretch of DNA. There is a single case, described in association with a pathogenicity island, where an integrase is able to catalyze integration and excision reactions, independent of a specific excisionase (30, 31), but this has never been described for an infective phage.

Preliminary qualitative RT–PCR data revealed that Stx-phage Φ24B weakly expresses a putative antirepressor in uninduced lysogen cells (20). Moreover, these cells also unexpectedly express int indicating that integration process may not be fully repressed in lysogen cells. These two observations support alternative hypotheses to explain establishment of multiple lysogeny by Stx-phages, antirepressor or integrase mediated, respectively. Here, we demonstrate increasing levels of Φ24B integrase expression in multiple lysogen cultures and apparently negligible repression of int expression by CI. The Φ24B int transcription start site and promoter region were determined and found to differ from in silico predictions (20). We also confirm activity of this integrase in the integration reaction but rule out the possibility that it can catalyze excision in the absence of a phage directionality factor, despite the absence of an obvious associated excisionase in the Φ24B genome. Subsequently, we located a novel excisionase gene, identified through genetic complementation assays when co-expressed with Φ24B int. In addition, the previous proposal of a putative antirepressor affect on superinfection immunity is revisited and a new hypothesis put forward.

MATERIALS AND METHODS

Bacterial strains, bacteriophages and media

The bacterial strains, bacteriophages and plasmids used in this study are listed in Table 1. Escherichia coli K-12 strain MC1061 was the host of choice for bacteriophage infection and lysogen production. JM109-λ pir was the propagation strain for the suicide plasmid pKNG101, while Invitrogen One Shot TOP10 cells were used for all cloning and recombination assays, except where the plasmid pΦ24B::IX was used, which required the use of E. coli strain BL21(AI) to control the expression of excisionase under its inducible T7 promoter. Bacteria were routinely cultured in phage buffer [2.5% (w/v) Luria–Bertani (LB) broth (Merk) with 0.01 M CaCl2] or on LB agar plates [2.5% LB broth plus 1.5% (w/v) agar (Merk)]. Where appropriate, the following antibiotics were added (μg ml−1): rifampicin (3500), ampicillin (100), kanamycin (50), chloramphenicol (50) and spectinomycin (100).

Antirepressor knockout

The ‘putative’ Φ24B antirepressor was amplified using Ant F/R primers (Supplementary Table S1). The resulting PCR product and the suicide vector pKNG101 were cut with BamHI restriction endonucleases, agarose gel purified and recovered with the QIAquick gel extraction kit. The two purified DNA fragments were ligated together to produce pKNG-ant (Table 1). Selected transformants were used as a template for inverse PCR using Ant Inv F/R primers (Supplementary Table S1), while the aadA spectinomycin resistance cassette was also amplified with aadA F/R AccI (Supplementary Table S1). Both PCR products were purified and cut with the restriction endonuclease AccI and ligated together to produce pKNG-antΔ (Table 1). pKNG-antΔ was transformed into electrocompetent MC1061 lysogen cells carrying either of the prophages Φ24B::Kan and Φ24B::Cat (Table 1). All SmR resistant transformant colonies contained pKNG-antΔ integrated into Φ24B ant, as confirmed by the absence of a PCR amplification product with Ant F/R primers and sensitivity to sucrose when subcultured on LB Agar plates in which NaCl was...
Table 1. Bacteria/phage strains and plasmids used in this study

<table>
<thead>
<tr>
<th>Strain/Plasmid</th>
<th>Description</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC1061</td>
<td>E. coli K-12 derivative</td>
<td>(83)</td>
</tr>
<tr>
<td>DM1187</td>
<td>E. coli K-12 derivative, recA441</td>
<td>(84)</td>
</tr>
<tr>
<td>JM109 λ pir</td>
<td>E. coli K-12 derivative, λ pir</td>
<td>(85)</td>
</tr>
<tr>
<td>pCDF-Duet</td>
<td>pCDF-Duet</td>
<td>(18)</td>
</tr>
<tr>
<td>pStxACYC</td>
<td>pStxACYC</td>
<td>(20)</td>
</tr>
<tr>
<td>pKNG101</td>
<td>pKNG101 R6K Ori, Sp/R Sm/R Novagen, EMD</td>
<td></td>
</tr>
<tr>
<td>pACYC-Duet</td>
<td>pACYC-Duet plasmid backbone, and the wild-type ant was replaced by ant ΔaadA, as confirmed by PCR (data not shown).</td>
<td></td>
</tr>
<tr>
<td>OneShot TOP10</td>
<td>Competent cells, recΔ araB araD::T7RNAP tetA, recA araBAD C mycHis6 epitope, pBR322 ori, Am^R</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>pKNG101</td>
<td>R6K Ori, sacB, Sm^R</td>
<td>(86)</td>
</tr>
<tr>
<td>pACYC-Duet</td>
<td>P15A ori, Cm^R Novagen, EMD</td>
<td></td>
</tr>
<tr>
<td>pCDF-Duet</td>
<td>CDF ori, Sp/Sm^R</td>
<td></td>
</tr>
<tr>
<td>pBAD/myc-His C</td>
<td>pBAD carrying myc Epitope, pKNG101 carrying Myc, Sm^R (86)</td>
<td></td>
</tr>
<tr>
<td>pKNG101</td>
<td>pKNG101 carrying int</td>
<td></td>
</tr>
<tr>
<td>pBAD/myc-His C</td>
<td>pBAD carrying myc Epitope</td>
<td></td>
</tr>
<tr>
<td>pKNG101</td>
<td>pKNG101 carrying pBAD carrying myc Epitope</td>
<td></td>
</tr>
<tr>
<td>pBAD/myc-His C</td>
<td>pBAD carrying myc Epitope</td>
<td></td>
</tr>
<tr>
<td>pKNG101</td>
<td>pKNG101 carrying pBAD carrying myc Epitope</td>
<td></td>
</tr>
<tr>
<td>pStxACYC</td>
<td>pStxACYC carrying myc Epitope</td>
<td></td>
</tr>
</tbody>
</table>

ant: putative antirepressor; int: integrase; cl: phase repressor; attB: phage insertion site within the MC1061 chromosome utilized by Φ24g attP: Φ24g phage attachment site; xis: excisionase; stxA: Shiga toxin A sub-unit for Stx2; aph^R: kanamycin resistance gene (Kan^R); adaA: spectinomycin resistance gene (Sp^R/Sm^R); cat: chloramphenicol acetyl transferase gene (Cm^R); luxABCDE: Photorhabdus luminescens bioluminescence operon; recA^-: mutation in DNA repair recombinase gene; recA441: mutation in recA leading to constitutive activation; λ pir: chromosomal copy of the bacteriophage lambda replication initiation protein; R6K ori: λ pir dependent plasmid origin of replication; tetA: tetracycline resistance; araBAD: arabino-inducible promoter; T7RNAP: bacteriophage T7 RNA polymerase; His6: C-terminal hexa-histidine tag; C-myc: C-terminal myc human oncogene peptide epitope; sacB: levansucrase, a conditional lethal gene; ori: origin of replication; pR*: Φ24g late gene promoter.).

Luminescent reporter gene fusion construction

The Φ24g pR* promoter region was amplified from purified Φ24g DNA using pStx F/R oligonucleotides (Supplementary Table S1) and the luxABCDE operon was amplified from pSB2030 (32) using Lux F/R primers (Supplementary Table S1). Both the promoter and lux operon were cut with the relevant restriction enzymes (PstI, SpeI and XhoI), while the medium copy number plasmid pACYC-Duet (Table 1) was cut with PstI and XhoI. Each of the above was agarose gel purified and recovered using the QIAquick gel extraction kit (Qiagen, Ltd). A single plasmid, pStxACYC, was constructed following ligation of the three products, (Table 1). pStxACYC was then transformed into competent MC1061/Φ24g::Kan or MC1061/Φ24g::Kan antA lysogen cells (Table 1); transformants were selected for on agar plates containing chloramphenicol (50 µg·mL^-1^) and spectinomycin (100 µg·mL^-1^). Complementation of ant was performed by amplifying the Φ24g ant gene from purified bacteriophage Φ24g DNA using Ant F/R oligonucleotides (Supplementary Table S1). This product was agarose gel purified and cut with the restriction endonucleases Ncol and SalI, whilst the plasmid pBAD/myc-His was cut with the same enzymes (Table 1). Both products were ligated together to create the antirepressor expression construct, pΦ24g-ant (Table 1), pΦ24g-ant was then transformed into competent Φ24g::Kan antA lysogen cells harboring pStxACYC.

Lux reporter gene assay

Fresh overnight cultures were used to inoculate 50 mL of LB broth; each culture was prepared in duplicate—one culture serving as a negative control and the other subjected to norfloxacin-based bacteriophage induction (33). Cultures were grown for ~3 h to mid-exponential growth phase (OD~600 = 0.5), at which point the antibiotic norfloxacin was added to the induced cultures at a working concentration of 1 µg·mL^-1^. Triplicate 200 µL samples were taken immediately for luminescence and optical density measurements, in opaque and transparent 96-well plates, respectively, and analyzed in a VICTOR3 plate reader (PerkinElmer). This sampling regime was repeated at intervals of 30 min. Specific luminescence was calculated as luminescence per optical density unit.

Quantitative PCR

Single, double and triple lysogen cultures were harvested at an OD~600~ of 0.5. Total RNA was extracted using an RNeasy mini kit (Qiagen Ltd), according to the manufacturer’s specifications. Genomic DNA contamination was removed with Ambion RNase-free DNase (Applied Biosystems) and RNA was quantified with a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific Inc.). RNA (0.5 µg) for each reaction was transcribed into cDNA using RNaseH reverse transcriptase (Bioline Ltd) in a total of 20 µL cDNA (1 µL) was then used as template for each qPCR reaction in Sensimix qPCR reaction mixture (Quantace Ltd). All experiments were designed to comply with the standard Sensimix cycling conditions with an annealing temperature of 60°C and an amplicon size of ~100 bp. qPCR reactions were carried out using a MiniOpticon thermal cycler (BioRad Labs Inc.). Integrase was amplified using int qPCR F/R, and the copy number was determined on a standard curve (copy number range = 5 × 10^2 to 5 × 10^8), providing absolute quantification of the amplified products.
CI overexpression

Φ24g::cI was amplified from bacteriophage Φ24g DNA purified according to the method of Rooks et al. (34) using CI F/R oligonucleotides (Supplementary Table S1), agarose gel purified and cut with NcoI and SalI restriction endonucleases; the plasmid pBAD/Myc-His was also cut with NcoI and SalI (Table 1). These products were ligated together to create pΦ24g::cI (Table 1). pΦ24g::cI was then transformed into competent E. coli lysogen cells [strain MC1061 carrying the Φ24g::Kan prophage (MC1061/Φ24g::Kan)], and transformants were selected for by resistance to the antibiotics kanamycin (50 μg mL⁻¹) and ampicillin (100 μg mL⁻¹). A MC1061/Φ24g::Kan culture carrying pΦ24g::cI was incubated at 37°C until it reached an OD₆₀₀ of 0.5, and cI expression was induced through addition of arabinose to working concentrations of 0.2% (w/v), 0.002% (w/v) or 0% (negative control), before further incubation at 37°C for 2 h. Samples were taken post-induction and processed for qPCR, as above, and for SDS–PAGE analysis. Transcripts of the cI, Q and int genes were amplified with cI qPCR F/R, Q qPCR F/R and int qPCR F/R, respectively (Supplementary Table S1); an endogenous reference, 16S rRNA processing gene rimM, was amplified with 16S qPCR F/R (Supplementary Table S1). Expression from test strains was quantified by the 2⁻ΔΔCₜ relative abundance method (35).

Rapid amplification of 5'-complementary DNA ends (5'-RACE)

The lysogen MC1061/Φ24g::Kan was cultured at 37°C to an OD₆₀₀ of 0.5. Total RNA was extracted using an RNaseasy mini kit, and immediately processed with a 2nd generation 5'/3'-RACE kit (Roche Applied Science) according to manufacturer’s instructions for 5'-amplification. DNA purification steps were carried out using the manufacturer’s recommended High Pure PCR Product Purification Kit (Roche Applied Science), and the custom oligonucleotide primers Int RACE Sp3, Sp2 and Sp1 (Supplementary Table S1) were used in conjunction with primers included in the 5'/3'-RACE kit for amplification of the 5'-integrase region. The resultant PCR amplification products were cloned into a ZeroBlunt vector (Invitrogen) and sequenced using standard M13 F/R primers (GATC-Biotech). The transcription start site was mapped onto the Φ24g genome, and the promoter region and binding sites predicted in silico.

Integration/excision vector construction

The Φ24g int and attP sites along with its primary bacterial attB site were amplified with Integrase F/R, attP F/R and attB F/R, respectively (20). The integrase gene and the two attachment sites were each cloned, separately, into ZeroBlunt cloning vector (Invitrogen), according to manufacturer’s guidelines. The cloned fragments were subsequently excised using NcoI and KpnI, or BamHI and XhoI endonucleases for the integrase gene or both the attachment sites, respectively. Three plasmids with compatible origins of replication were also cut, pBAD/Myc-His C (NcoI and KpnI), pACYC-Duet and pCDF-Duet (BamHI and XhoI) (Table 1). After deactivation of the enzymes, the integrase gene was ligated into pBAD/Myc-His C to form pΦ24g-int (Table 1), attB was ligated into pACYC-Duet to form pΦ24g-attB (Table 1) and attP was ligated into pCDF-Duet to form pΦ24g-attP (Table 1).

Integration/excision assay

Various combinations of plasmid constructs in appropriate strains were produced as summarized in Table 2. Integration of pΦ24B-attP into pΦ24B-attLR to create pΦ24B-attL.R (Table 1) under different conditions was monitored qualitatively by plasmid purification using a Plasmid Miniprep kit (Qiagen Ltd). Cultures 2 and 4 were induced with arabinose, whilst 5 was induced with arabinose [0, 0.2 and 0.02 (w/v)] and IPTG (1 mM), each for 2 h, before plasmids were purified. Culture 7 was induced to produce native excisionase by addition of 1 μg ml⁻¹ norfloraxcin for 2 h, prior to plasmid purification.

Excisionase cloning and expression

The putative excisionase was initially identified by a systematic in silico screening process, based solely on size and location of possible ORFs. All potential ORFs of 150–600 bp, within 5 kb of Φ24g int and beginning with any of the E. coli alternate start codons [ATG, GTG, TTG, ATT and CTG (36)], were analyzed. The candidate Xis

<table>
<thead>
<tr>
<th>Table 2. Integration/excision plasmid permutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assay</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
</tbody>
</table>

A summary of the various combinations of plasmids used for recombination assays to assess the function of Φ24g integrase and excisionase. All of the assays summarised above were carried out in recA⁻ E. coli hosts. Plasmid components (pΦ24g-attB, pΦ24g-attP, pΦ24g-attLR, pΦ24g-int and pΦ24g-attP 1X) are described in Table 1, whilst the final column refers to presence/absence of a Φ24g prophage. For all of the assays, cells carrying the plasmids detailed were grown to mid-logarithmic growth phase (OD₆₀₀ 0.5) with the appropriate antibiotic selection. Plasmids from 1, 3 and 6 were immediately purified with a Plasmid Miniprep kit (Qiagen Ltd). Cultures 2 and 4 were induced with arabinose, whilst 5 was induced with arabinose [0, 0.2 and 0.02 (w/v)] and IPTG (1 mM), each for 2 h, before plasmids were purified. Culture 7 was induced to produce native excisionase by addition of 1 μg ml⁻¹ norfloraxcin for 2 h, prior to plasmid purification.
was amplified from Φ24B DNA using Xis F/R oligonucleotides (Supplementary Table S1), purified and cut with the restriction endonucleases NcoI and BamHI. pACYC-Duet (Table 1) was also cut with the same enzymes, NcoI and BamHI. The two compatible DNA fragments were ligated together to produce pΦ24B-X (Table 1). pΦ24B-X was then used as a template for PCR amplification with T7 F/Xis R. The resulting PCR product, containing the T7 promoter from pACYC-Duet plus the putative Φ24B xis, was cut with HindIII endonuclease and ligated into HindIII cut pΦ24B-int (Table 1) to create pΦ24P-IX (Table 1). pΦ24P-IX was transformed into BL21(AI) cells (Table 1) and transformants selected for by resistance to ampicillin (100 μg ml⁻¹). Excisionase activity was monitored by agarose gel electrophoresis and qPCR (see Integration/excision assay).

Bioinformatics

Database searching was done with BLAST and PSI-BLAST (37) in the nr database (38). Domain searches were made in the databases CDD (38) and InterPro (39). More sensitive domain searching was done with HHpred (40,41). In accord with the extreme sequence divergence exhibited by the Xis sequences, a structure-based alignment approach was employed. A structure alignment of members of Pfam (42) families Phage_AlpA [PF05930: TorI with PDB (43) code 1zh4 (44)], Excisionase like [PF07825: lambda phage Xis, 2ief (27)] and PF09035 [Tn916-Xis; 1y6u (45)] was made with MUSTANG and STACCATO (46) and used, as the basis to unite three alignments using the program Jalview (47). Each of the three alignments, made with MUSCLE (48), contained Xis sequences bearing similarity to the respective Pfam families. Other Xis sequences, belonging to different Pfam families but recognizably distantly related to AlpA, were added to the aggregate alignment with the benefit of alignments deriving from the sensitive profile–profile matching program HHpred (40,41). The resulting alignment was subject to phylogenetic analysis with MEGA 4 as follows (49). The bootstrap consensus tree inferred from 1000 replicates was taken to represent the evolutionary history of the taxa analyzed (50). The evolutionary distances were computed using the Poisson correction method (51) and the units are the number of amino acid substitutions per site. The ME tree was searched using the Close-Neighbor-Interchange (CNI) algorithm (52) at a search level of 1. The Neighbor-Joining algorithm (53) was used to generate the initial tree. All positions containing alignment gaps and missing data were eliminated only in pair-wise sequence comparisons.

Ten models of Φ24B Xis were produced with MODELLER (54) using the structure of TorI [PDB code 1zh4 (44)] as a template. The DOPE statistical potential (55) and stereochemical analysis (56) were used to select a model for further analysis. Electrostatic characteristics of the model were explored with APBS (57) and putative protein–protein interfaces sought with HotPatch (58). DNA-binding ability was tested with DB-MOM (59) and PRO-DNA (60). Further protein structure superpositions were done with SSM (61) and displayed in PyMOL (62), which was also used to produce structure figures.

RESULTS

Characterization of the activity of ant

Previous studies have shown that the frequency of multiple lysogen formation, through sequential superinfection, exceeds the rate of single lysogen formation by infection of a naïve host (19,20). This suggests that the incumbent prophage in some way aids the integration of the superinfecting bacteriophage, thus increasing integration efficiency. The presence (Figure 1A) and expression of a putative antirepressor gene (ant) present in Φ24B was postulated to be responsible for the phenomenon of multiple lysogeny through a mechanism of derepression (20). However, superinfection experiments using Φ24B antA strains demonstrated that the lack of Ant failed to have any effect, positive or negative, on superinfection compared with wild-type Φ24B (data not shown). Instead, further investigation has produced evidence of a likely role in assisting relief of CI repression during induction into the lytic cycle, consistent with independent observations reported for a similar antirepressor of bacteriophage 2851 (63). A Pφ'luxABCDE fusion was used as a reporter for bacteriophage late gene expression following norfloxacin induction of lysis (Figure 1B). Three different lysogen cultures were tested: Φ24B wild-type, Φ24B antA and Φ24B antA complemented with pBAD-ant. Luminescence was measured over 4 h proceeding norfloxacin addition. The lag phase before substantial increases in Pφ' expression were detectable was ~120 min for each of the strains, and the difference between both the two test strains and the wild-type reference varied markedly from this time point onwards (Figure 1C). The antA strain peaked at luminescence intensity up to 55.6% of the wild-type luminescence at 120 min, but decreased to 27.2–31.0% for the remainder of the experiment (Figure 1C). In comparison, the complemented strain was able to restore 82.1% of wild-type luminescence levels at 120 min before decreasing to 64.3–65.3% and eventually returning to 79.5% after 240 min (Figure 1C).

Uninduced integrase expression

Clearly, some factor outside the production of an antirepressor is involved in driving increased frequencies of lysogen formation upon superinfection. A potential mechanism to explain this observation would be expression of integrase from the resident integrated phage, which would more typically be strongly repressed by the lambdoid immunity system present in Stx-phages. Integrase transcript abundances were determined for single, double and triple lysogens of the Stx-phage by qPCR, in the absence of any stimulus for the lytic cycle. Expression of Φ24B integrase increased 2.6-fold in the double lysogen and 7.5-fold in the triple lysogen, each compared to expression levels in the single lysogen (Figure 2A). This observation suggests that integrase expression is not fully controlled by the global phage...
repressor protein, CI. Furthermore, complete genome sequence data for \( \Phi C8 \) (Accession number: HM_208303) has revealed that the integrase gene is present in an inverted orientation (Figure 3A) compared to other lambdoid phages such as lambda (NC_001416) and 933W (NC_000924). This inversion implies that its regulation must be different to these well-studied phages like Lambda phage where the \( \text{int} \) gene is subject to repression from CI. To identify whether the regulation of \( \Phi C8 \) Int expression is uncoupled from CI repression, CI was overexpressed through arabinose induction (Figure 2B) in a \( \Phi C8 \) lysogen background, and the effect on integrase expression monitored by qPCR (Figure 2C). As a control, expression of the \( Q \) antiterminator gene, the key regulator of the lytic cycle, was also monitored (Figure 2C). As expected, integrase expression was not affected by increased levels of CI, whereas \( Q \) expression was significantly \((P < 0.05)\) decreased (Figure 2B and C). The \( \Phi C8 \) immunity region, whilst identical to bacteriophage 933W (20,64), differs in organization from bacteriophage lambda in that there are only two left operator binding sites. (A) Transcription is indicated by arrows above or below the line with \( pR \), \( pR' \) and \( pRM \) promoters labeled adjacent to their transcription start sites. Gene names, indicated in italics, are transcribed leftward when below the line or rightward when above. The right operator region including the 3 CI binding sites (OR1-3) is annotated. The qut transcription termination site is represented by a raised triangle immediately after \( pR' \). The putative antirepressor gene is located before this qut site and thus should be expressed shortly after induction to the lytic cycle, where it plays a role in accelerating phage derepression. Relief of CI mediated repression results in high level expression from \( pR' \) leading irreversibly to lysis and release of \( S\text{tx} \). (B) Map of the \( pR' \) reporter gene fusion, including the \( s\text{tx}A \) transcription start site. The \( pR' \) promoter and the \( qut \) transcription termination site are indicated. (C) Lux specific luminescence driven from the \( pR' \) promoter during induction of phage lysis by addition of norfloxacin antibiotic. \( T_0 \) is measured from the addition of inducing agent and there is a 90-120min lag phase before expression increases appreciably. The three lines correspond to the antirepressor knock-out mutant (\( \Delta \), Ant\( - \)), the complemented strain (\( \Box \), Ant\( + \)) and the wild-type control (\( \square \), WT). Error bars represent combined standard error for luminescence/optical density quotient \((n = 3)\).
regions rather than three. This does not however compromise repression of transcription from the pL promoter or affect lysogen stability (65,66). Bandshift analysis was carried out to confirm that the CI protein is able to effectively bind to the left operator region. Indeed, similar to 933W (67), binding occurred as expected (data not shown), and thus a defect in the immunity region is unlikely to be a factor in the phenomenon of multiple lysogeny.

5'-RACE determination of integrase start site

Φ24B encodes a novel integrase, as yet uncharacterized, which is extremely promiscuous with regard to integration site selection (20). Furthermore, there was no obvious gene encoding an excisionase-like protein identifiable anywhere on the bacteriophage chromosome (HM_208303), including the areas immediately up- and downstream of the int gene. An integrase capable of performing both excision and integration functions has been described in association with an E. coli pathogenicity island (31), but never in association with an inducible bacteriophage, lambdoid or otherwise. In silico analysis predicted a 1245 bp integrase-encoding ORF in the Φ24B genome, and this was cloned into a pBAD expression vector (Invitrogen). However, we failed to successfully express the integrase protein after repeated attempts (data not shown). We therefore set out to establish the transcription and translation start sites, and determine whether these were compatible with the in silico predictions. To resolve the transcriptional start site, rapid amplification of the 5'-complementary DNA end (5'-RACE, Roche) was performed. The cDNA end product was sequenced and revealed that transcription actually initiates 46-bp downstream of the expected translational start codon (Figure 3B). From this information, the promoter elements were predicted using the BDGP neural network promoter predictor (68) and a potential UP-transcription activation element was mapped by comparison to published sequences (69,70). UP elements are short stretches of AT-rich sequence present in some promoters, directly upstream of the core region. These elements enhance transcription by interacting with RNA polymerase, though they are not essential for overall promoter function (69–71). The Φ24B integrase promoter (pInt) contains a match to only the distal section of the UP element consensus sequence (9 out of 11 bases) and the proximal section is absent, however, ‘half sites’ have previously been shown to be functional (72). Improved promoter and transcriptional start data also assisted in the production of a new int ORF prediction using the FrameD online ‘Gene Finder’ tool (http://bioinfo.genotoul.fr/apps/FrameD/FD) (73). This updated ORF utilizes the rare TTG start codon and experimentally facilitated Int expression from an inducible promoter (Figure 3C) (73).

Integration excision system: integrase inability to catalyze excision

In order to test Int function experimentally, a three-plasmid system was used in which the Φ24B attP site, bacterial attB site and the Φ24B integrase were each cloned separately into compatible, low copy number

Figure 3. (A) Schematic map of the Φ24B integrase region. This updated genetic map includes the integrase in an inverted orientation to the bacteriophage lambda model, and thus not under the control of CI through its indirect repression of int transcription from pL during lysogeny (as in lambda), and the novel excisionase ORF which would be under the auspices of CI via pL. The proposed pInt promoter is also indicated. (B) Schematic of the Φ24B integrase transcription start site and predicted promoter as determined by 5'-RACE and in silico analysis. The poly-T start to the RACE output (5'-RACE) is an artifact of the amplification process; Φ24B genome sequence (HM_208303) was used for the alignment. Putative -10/-35 promoter constituents and the distal portion of a putative UP-element are indicated by labeled solid lines and a dashed line, respectively. Transcription start is also indicated by asterisks along with the new putative translational start site hash and ribosomal binding site (RBS) along with the previous, discredited translational start site (X). (C) Expression of Φ24B integrase from an araBAD inducible promoter with increasing levels of arabinose, shown by 10% SDS–PAGE.
plasmids. Integration and excision events could then be detected by determining the presence of either a large \( p\Phi 24\) integrated plasmid or individual, resolved \( p\Phi 24\)-attB/attP containing plasmids (Figure 4A), and these could be quantitated further by qPCR. By this method, we were able to demonstrate integrative function for the expressed integrase protein alone, in the absence of any other accessory phage proteins (Table 2 and Figure 4B). Moreover, efficient integration of the plasmids was also observed when transformed into a single lysogen host in the absence of any stimulus for integrase production (Figure 4B). This latter observation further corroborates the hypothesis set out earlier that integrase is expressed in a lysogen, uncoupled from CI mediated repression, and that this may help to assist in the integration of superinfecting bacteriophage. All attempts to resolve the integrated \( p\Phi 24\)-attLR plasmid into its two component plasmids with the integrase expression construct present, alone, were futile (Table 2, assay 4; data not shown), whilst resolution of \( p\Phi 24\)-attLR within a lysogen was easily detectable when that lysogen was subjected to norfloxacin induction (data not shown), but not detectable, even by qPCR, in an uninduced lysogen (Table 2, assay 7). These experiments indicated that there must be a phage-encoded accessory factor that remained coupled to CI repression and the standard lambdoid regulatory cascade that is essential for \( 
abla \Phi 24\) prophage excision.

Identification and bioinformatic analysis of a putative excisionase

The small size of the majority of currently described excisionases (typically more than 100 amino acids) limits
the ability to predict and identify novel excisionase proteins based on homology searches (28). Excisionase genes are usually found in close proximity to their cognate int (74,75), therefore, due to the location and orientation of \( \Phi 24_b \) int, it was predicted that the excisionase gene was likely to be situated adjacent to the 3’ of the integrase. In addition, although \( E. coli \) primarily utilizes ATG as a translation start codon (83% of genes in \( E. coli \) K-12), several other codons can also be used such as GTG (14%) and TTG (3%) as well as the rare ATT and CTG codons (36). Bearing this information in mind, investigation of all possible short ORFs in the proximity of the integrase gene identified only one candidate ORF, which was predicted to encode an 8 kDa protein. Database searches by BLAST and PSI-BLAST revealed only four close homologs from \( E. coli \) strains, none of which had been experimentally determined to possess an excisionase function. Similarly, searches of domain databases CDD and InterPro with default parameters failed to place the candidate protein sequence in any known protein family. Nevertheless, analysis with HHpred, a more sensitive sequence comparison method, revealed a clear but distant relationship between the candidate ORF and Pfam family phage_AlpA (PF05930). The e-value was \( 5 \times 10^{-6} \), and the alignment between the ORF and the domain consensus produced 40% sequence identity, although the alignment only covered 47 residues of the candidate ORF. This provided a strong indication that the candidate ORF indeed encoded a Xis protein since several characterized Xis/RDF proteins belong to this family e.g. \( E. coli \) TorI (76), and the Xis proteins from \( E. coli \) phage Phi80 (77) and Shigella flexneri (78).

An alignment was made of \( \Phi 24_b \) Xis with other experimentally demonstrated or reliably predicted excisionase sequences, building on earlier lists (28,78). Extreme sequence diversity has often left doubts over the true homology of some excisionases (28), and the alignment was therefore limited to sequences with demonstrable homology, using sensitive profile–profile matching (40,41) to winged-helix excisionases such as lambda Xis and TorI. As well as phage_AlpA, these sequences belonged to, or resembled, several Pfam families: Excisionase like [PF07825; e.g. lambda phage Xis (27)], Tn916-Xis [PF09035 (45)], Putative excisionase [PF06806 (64)] and Regulatory phage protein Cox [PF10743 (79)]. Alignment (Supplementary Figure S1) of these sequences illustrated the extreme sequence diversity of excisionases with this architecture. Even within the common helix-turn-helix motif and the shared small ‘wing’ \( \beta \)-sheet sequence, conservation is very limited, and many members have N- or C-terminal extensions. This diversity is responsible for the difficulty previously encountered (28) and seen again in this work, of annotating excisionase proteins using conventional bioinformatic tools. Bootstrap phylogenetic analysis of the alignment (Supplementary Figure S2) revealed few strongly supported nodes, unsurprising in view of the diversity of the sequences. In particular, \( \Phi 24_b \) Xis clustered reliably only with an unpublished, unannotated sequence from Stx2-converting \( \Phi 86 \) phage (Accession number: AB255436) from \( E. coli \) O86. This phage is a close relative of \( \Phi 24_b \) and its almost identical integrase shares the inverse orientation, in contrast, to most characterized lambdoid phages.

**Experimental validation of excisionase function**

Excisionase function was confirmed experimentally for the newly identified \( \Phi 24_b \) xis ORF by co-expression with \( \Phi 24_b \) int from a single expression plasmid construct, \( \Phi 24_b \) intX, in the presence of the integrated \( \Phi 24_b \)-attLR plasmid in situ (Figure 4). The co-expressed proteins were able to effectively resolve \( \Phi 24_b \)-attLR into its constituent plasmids, resulting in cells harboring the expression construct plus the plasmids \( \Phi 24_b \)-attB and \( \Phi 24_b \)-attP, as observed by agarase gel electrophoresis (Figure 4B). Some residual \( \Phi 24_b \)-attLR remained in these cultures, though in lower quantities than the resolved plasmids. To obtain a quantitative measure of the efficiency of these reactions, an absolute qPCR method was devised to enumerate the levels of integrated \( \Phi 24_b \)-attLR plasmid compared to individual component plasmids (Figure 4A). Integration and excision reactions were consistently efficient and indisputably substantiate the claim that the proposed \( \Phi 24_b \) integrase and excisionase do indeed function as predicted (Figure 5).

**A molecular model of \( \Phi 24_b \) Xis**

The structure of \( E. coli \) TorI (44) was used as a template to model \( \Phi 24_b \) Xis. Although the HHpred-derived alignment matched only 15% of residues between the two proteins, modeling was facilitated by the existence of only a single one-residue deletion in \( \Phi 24_b \) Xis compared to TorI, which could be readily structurally accommodated. Ten models were constructed with MODELLER (54) and that with the best DOPE score (55) selected. The chosen model also had good stereochemical qualities including 87%
of residues in core areas of the Ramachandran plot (56). The C-terminal seven residues of \( \Phi 24_{B} \) Xis could not be modeled since they were not aligned with template by HHpred and are likely to adopt a different conformation: no regular secondary structure for this region is predicted in \( \Phi 24_{B} \) Xis, but the corresponding part of TorI is \( \alpha \)-helical.

DNA-binding proteins often have a large positively charged surface and we sought this feature on the \( \Phi 24_{B} \) Xis model. Indeed, one face of the Xis protein bears a pronounced positive charge (Figure 6A). This corresponds to the side of the protein structure that binds DNA in the lambda phage Xis crystal structure (Figure 6C) (27), which has been implicated by site-directed mutagenesis in DNA binding of TorI (44).

In view of the capacity of some Xis proteins—demonstrated or predicted (27)—to self-interact and form DNA-bound micronucleoprotein filaments, potential sites of protein–protein interaction on the \( \Phi 24_{B} \) Xis model were sought with HotPatch (58). For protein–protein interactions, exposed hydrophobic residues constitute the most reliable predictor of interaction sites. The
model of $\Phi 24_B$ Xis produces a single, strong prediction comprising four solvent-exposed hydrophobic residues in close proximity—Phe31, Pro32, Val35 and Leu36 (Figure 7).

DISCUSSION

Previous examination of the $\Phi 24_B$ genome sequence revealed the presence of a putative antirepressor gene possessing 37% amino acid identity to the well-characterized P22 antirepressor, Ant (Accession number: NP_059643) (20,25). It was also demonstrated qualitatively that the ant gene was expressed at very low levels in a quiescent lysogen culture (20). Alignment of the putative $\Phi 24_B$ antirepressor protein with its P22 counterpart revealed significant sequence similarity primarily in the C-terminal domains (82% identity for the final 104 C-terminal residues). When expressed, P22 Ant is able to inactivate the phage repressor protein, analogous to the lambda CI repressor. This observation and the fact that we could measure expression of ant in the lysogen (20) led to speculation that $\Phi 24_B$ Ant could alleviate CI repression sufficiently to allow superinfecting phages to integrate. However, the data presented here refute this hypothesis. No significant effect was observed on lysogen formation rates, either when superinfecting an isogenic ant-strain or when Ant was expressed from a plasmid in the host cells undergoing superinfection. On the other hand, when an ant-lysogen strain was induced into the lytic life cycle it was unable to effectively upregulate expression from the ant underlyng superinfection. On the other hand, when Ant was expressed from a plasmid in the host cells (Figure 2) nor coupled to the lambdoid phage regulatory networks, but rather its expression is allowed to continue during stable lysogeny (Figure 2), presumably leading to an intracellular build up of Int. It is the accumulation of integrase in the cytoplasm, in conjunction with the lack of repression of incoming isogenic phage integrase production, which most likely accounts for the increased frequency of lysogen formation during superinfection compared to infection of a naïve host (20).

Furthermore, as the excisionase is not produced from the $pInt_{24_B}$ promoter, $\Phi 24_B$ is able to maintain a state of stable lysogeny despite the constitutive expression of int, as xis expression is linked to a standard lambda prophage induction pathway. This is exemplified by the almost complete conversion of plasmids bearing $attP$ and $attB$ to the integrated form (Figures 4 and 5), when propagated in an uninduced $\Phi 24_B$ lysogen. In summary, the evidence presented points to a clear mechanism for $\Phi 24_B$ multiple lysogen production at a high rate, due to the fact that integrase is outside the control of global phage repression, however, prophage stability is maintained due to the orientation and lambdoid genetic regulatory strategy which leaves xis coupled to CI repressor control.

Identification of the $\Phi 24_B$ excisionase was a non-trivial task as, despite careful investigation of the $\Phi 24_B$ genome sequence, we were unable to identify any candidate ORFs likely to encode an excisionase gene in the expected area, and Blastn searches of the sequences surrounding the int gene against the non-redundant databases failed to elucidate any possible Xis coding sequences. This led us to consider the possibility that there may be no such gene present in the phage genome. Although there have been no reported cases of infective bacteriophages able to catalyze integration and excision with an integrase alone, such a situation has been described for a uropathogenic E. coli pathogenicity island (30,31). However, in light of the fact that resolution of p$\Phi 24_B$-attLR to its constituent plasmids was not possible when integrase alone was expressed from an inducible plasmid while excision events were detected when the $\Phi 24_B$ prophage was present, a systematic location-targeted strategy was successfully employed. Distant homology of the candidate Xis protein sequence with the Phage_AlpA family, already containing several known excisionases, bore out our initial suspicions based on its genomic context. A molecular model confirmed that $\Phi 24_B$ Xis protein adopts a winged-helix fold.

Interestingly, excisionase proteins with the winged-helix fold appear to function in diverse ways. In the archetypal Lambda phage Xis, from Pfam family ‘Excisionase-like’, binds directly to DNA, forming a micronucleoprotein filament which bends the DNA duplex in a way predicted to facilitate the excision process (27). Data suggest that the Tn916 Xis may function similarly (27). Escherichia coli Tor1, from Pfam family phage_AlpA acts as a response regulator inhibitor, binding not only to DNA (44) but also to the DNA-binding domain of the TorR protein (76). Although this protein–protein interaction does not affect the DNA-binding capacity of TorR, it is thought to impede recruitment of RNA polymerase to the torC promoter. Finally, the AlpA protein acts as a transcription factor for the int gene of the phage CP4-57 in E. coli K-12.
(81,82), thereby affecting excision by a third variant mechanism.

DNA binding is a common feature of all these excisionases, and so it is no surprise that the model of \( \Phi 24b \) Xis bears a characteristic positively charged surface (Figure 6). Structure-based predictors of DNA-binding capacity—DB-MOM (59) and PRO-DNA (60)—confirm that \( \Phi 24b \) Xis should bind DNA. Attempts to predict further details of the molecular mechanism of \( \Phi 24b \) Xis by phylogenomic analysis failed since bootstrapped trees did not reliably group this Xis with other experimentally studied excisionases (Supplementary Figure S2). However, the \( \Phi 24b \) Xis model allowed us to address one aspect of function, namely the ability to form micronucleoprotein filaments. Inspection of the structure of phage lambda micronucleoprotein filament structure (Figure 7A) highlights the role of the loop containing Arg13 and Arg14 in forming inter-subunit contacts. The corresponding loops in \( \Phi 24b \) Xis (Figure 7B) and Tn916 Xis (not shown) are shorter and, without significant rearrangement, appear to be too short to form similar interactions. However, seven residues at the C-terminus of \( \Phi 24b \) Xis could not be modeled since they did not align reliably with the template. Intriguingly, the C-terminus of the incomplete model clearly shows that the missing residues could interact with the neighboring subunit at exactly the region picked out as a potential protein–protein interaction site. DNA purified from cells expressing \( \Phi 24b \) Xis could not be cut with restriction endonucleases without a significant degradation of DNA; no bands were seen on subsequent agarose gels, only smearing (Figure 4B). This observation offers circumstantial support for the notion that \( \Phi 24b \) Xis binds DNA as micronucleoprotein filaments thereby destabilizing the duplex structure and increasing its susceptibility to degradation.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

Funding for open access charge: Funding for the work preceding this article was obtained from the Biotechnology and Biological Science Research Council of the U.K. The work detailed in this article was funded by its authors.

Conflict of interest statement. None declared.

REFERENCES


