Mobile Element Integration Reveals a Chromosome Dimer Resolution System in Legionellales

Beth Nicholson,a Shayna R. Deecker,a Alexander W. Ensminger*a,b

aDepartment of Biochemistry, University of Toronto, Toronto, Ontario, Canada
bDepartment of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada

ABSTRACT In bacteria, the mechanisms used to repair DNA lesions during genome replication include homologous recombination between sister chromosomes. This can lead to the formation of chromosome dimers if an odd number of crossover events occurs. The dimers must be resolved before cell separation to ensure genomic stability and cell viability. Dimer resolution is achieved by the broadly conserved dif/Xer system, which catalyzes one additional crossover event immediately prior to cell separation. While dif/Xer systems have been characterized or predicted in the vast majority of proteobacteria, no homologs to dif or xer have been identified in the order Legionellales. Here, we report the discovery of a distinct single-recombinase dif/Xer system in the intracellular pathogen Legionella pneumophila. The dif site was uncovered by our analysis of Legionella mobile element-1 (LME-1), which harbors a dif site mimic and integrates into the L. pneumophila genome via site-specific recombination. We demonstrate that lpg1867 (here named xerL) encodes a tyrosine recombinase that is necessary and sufficient for catalyzing recombination at the dif site and that deletion of dif or xerL causes filamentation along with extracellular and intracellular growth defects. We show that the dif/XerL system is present throughout Legionellales and that Coxiella burnetii XerL and its cognate dif site can functionally substitute for the native system in L. pneumophila. Finally, we describe an unexpected link between C. burnetii dif/Xer and the maintenance of its virulence plasmids.

IMPORTANCE The maintenance of circular chromosomes depends on the ability to resolve aberrant chromosome dimers after they form. In most proteobacteria, broadly conserved Xer recombinases catalyze single crossovers at short, species-specific dif sites located near the replication terminus. Chromosomal dimerization leads to the formation of two copies of dif within the same molecule, leading to rapid site-specific recombination and conversion back into chromosome monomers. The apparent absence of chromosome dimer resolution mechanisms in Legionellales has been a mystery to date. By studying a phage-like mobile genetic element, LME-1, we have identified a previously unknown single-recombinase dif/Xer system that is not only widespread across Legionellales but whose activity is linked to virulence in two important human pathogens.

KEYWORDS Legionella pneumophila, Coxiella burnetii, chromosome dimer resolution, site-specific recombination, Xer, dif, IMEX, Coxiella, Legionella, bacteriophages, mobile genetic elements

The circular nature of bacterial chromosomes brings about a problem for the partitioning of genetic information to daughter cells: the formation of chromosome dimers during replication. Dimers are generated when homologous recombination-mediated repair of DNA lesions results in an odd number of crossover events between sister chromosomes. The chromosome dimers must be resolved into monomers for proper cell division to occur. The important and broadly conserved bacterial dif/Xer system overcomes this topological problem by catalyzing one additional crossover...
event immediately prior to cell separation (1, 2). The dif (deletion-induced filamentation) site is a conserved ~30-bp DNA element on the chromosome that contains two binding sites for tyrosine recombinases—either two copies of the same recombinase or more commonly two different recombinases (3). In a highly regulated process, a nucleoprotein complex containing the Xer tetramer and two aligned dif sites forms a recombination synapton, and site-specific recombination between them resolves the chromosome dimer into monomers (4–6).

Dimer resolution via site-specific recombination by tyrosine recombinases was first discovered in plasmids (7, 8), including CoIE1, in which a DNA sequence named cer was found to be essential for plasmid monomerization and stability (8). Recombination between cer sites was later found to require catalysis by XerC (9) and XerD (10) encoded on the Escherichia coli genome. These findings, along with homology between the chromosomal dif site and cer, led to the discovery of the dif/XerCD chromosome dimer resolution (CDR) pathway in E. coli (10–13) and later in several other species of bacteria (14–19) and archaea (20, 21). Subsequent dissection of the dif/Xer recombination machinery has been aided by another group of mobile genetic elements, named integrative mobile elements exploiting Xer (IMEXs) (22–27). IMEXs, which include lysogenic phages, genomic islands, and plasmids, contain a dif mimic sequence that allows them to integrate into the chromosome at the dif site using the host’s Xer recombination machinery (5, 28).

Homologs of xerC and xerD have been predicted to be present in the vast majority of proteobacterial species (29). Exceptions to this include two groups of bacteria that utilize a single-recombinase system. The streptococci/lactococci use a single recombinase, XerS, and an atypical dif site (16), and a group of Epsilonproteobacteria, including Helicobacter and Campylobacter species, use a single-recombinase XerH system (14, 17, 29). While dif/Xer homologs have been detected in almost 90% of proteobacteria, no dif or Xer homologs could be detected in the order Legionellales, despite encoding RecA, RecBCD, and RecF, which are thought to be responsible for dimer formation during replication (29). How then do the members of the order Legionellales overcome the threat to genome stability that chromosome dimers impose? It is possible that they use an entirely different method for CDR or that its dif/Xer system is sufficiently divergent from others to avoid detection by sequence-based homology searches.

Several strains of Legionella pneumophila harbor a phage-like integrative mobile genetic element named Legionella mobile element-1 (LME-1) (30–33). We previously found that integration into the genome requires a 22-bp attachment site (att) in LME-1 that is identical to a sequence on the L. pneumophila genome (33). Even though LME-1 is only integrated in a small proportion of sequenced L. pneumophila isolates, the chromosomal att site is present in all L. pneumophila isolates and Legionella species sequenced to date (see Table S1 in the supplemental material), a level of conservation that suggests an important function. We report here that the LME-1 att site on the chromosome meets several criteria of the missing dif site in Legionella: size of ~30 bp, low GC content, some degree of palindromicity, is positioned close to the replication terminus, and is in a noncoding region with varying flanking genes. We identify a single recombinase that is both necessary and sufficient for site-specific recombination at this site. Together, our results indicate that L. pneumophila uses a distinct single-recombinase system for CDR and that LME-1 is an IMEX of that system. We also report that xerL orthologs can be found across the Legionellales order. Finally, we show that the Coxiellaceae dif site and xerL ortholog can functionally substitute for the Legionella dif/XerL system and propose that the migration of Coxiella burnetii XerL off the chromosome has contributed to the stability of virulence plasmids in this pathogen.

RESULTS

The LME-1 attachment site is invariably close to the Legionella replication terminus. We previously determined that LME-1 integrates into the L. pneumophila genome via a 22-bp att site (33). The corresponding att site on the chromosome is contained within a broader 29-bp sequence that is conserved across all sequenced L.
pneumophila strains and Legionella species (Fig. 1A and see Table S1 in the supplemental material), suggesting that it performs an important function. While the att site itself is conserved, its genomic neighborhood varies (33) (Table S1), indicating that the intrinsic function of this sequence may be unrelated to its flanking genes. In a search for attributes of the att site that may hint at its function, we analyzed its chromosomal location in all L. pneumophila strains with available genome sequences (Table S1). We found that despite the disparate gene neighborhoods, the location of the att site was roughly opposite the origin of replication in all cases. In fact, GC skew analysis of each genome revealed that the att site was invariably close to the cumulative GC skew maximum, which occurs at the site of replication termination (34) (Fig. 1B and Table S1). The att site is present within a small ~8-kb window between 35 and 43 kb from the GC skew inflection point (Table S1). This specificity of positioning extended to the att sites of other Legionella species, which are located between 89 bp and 79 kb from the point of GC skew inflection (Fig. 1B and Table S1).

The proximity of the LME-1 att site to the terminus region is reminiscent of bacterial dif sites, so we next compared the att site to the dif sites of several bacteria with established dif/Xer systems. In addition to its proximity to the terminus, the LME-1 att site shows several similarities to known dif sites, including (i) intergenic location (Table S1), (ii) size of ~30 bp, (iii) low G+C content, and (iv) a short central region flanked by left and right arms with some degree of dyad symmetry (Fig. 1C). However, the LME-1 att site is distinguished from known dif sites by the perfect dyad symmetry of its 10-bp arms (Fig. 1C). Furthermore, these arms do not show homology to any known Xer-binding motifs (29).

The L. pneumophila dif site supports highly efficient RecA-independent DNA recombination. The similarities between the chromosomal LME-1 att site (here referred to as the Legionella dif site) and established dif sites suggest that it likely contains two binding sites for one or more tyrosine-type recombinases that catalyze site-specific recombination (3). To assess the ability of the L. pneumophila dif site to undergo site-specific recombination, we used an intermolecular recombination assay (Fig. 2A). In this assay, a nonreplicative plasmid (pJB4648) containing a gentamicin resistance marker and a 22-bp portion of the dif site (corresponding to the LME-1 att site) is transformed into L. pneumophila strain Lp02. If site-specific recombination occurs between the att site on the plasmid and the chromosomal dif site, the entire plasmid will be integrated into the genome and confer gentamicin resistance (Gmr) to the cell. We compared the number of transformants resulting from the att-containing plasmid to one with a control 22-bp sequence derived from the L. pneumophila genome near the dif site. To evaluate the efficiency of recombination supported by the dif site, we also included a plasmid containing 2,200 bp of unrelated intergenic sequence from L. pneumophila, which can integrate into the genome via a single-crossover homologous recombination event. Considering that chromosomal dif site-specific recombination is known to be RecA independent (13, 16, 35), we also assessed the integration of each of these plasmids in a recA deletion strain (Fig. 2B). To control for any strain-to-strain differences in overall transformation efficiency, we normalized the number of transformants for each plasmid to the number resulting from transformation with a replicative plasmid (pJB1806_GmR). The att site-containing plasmid integrated into the genome at a high efficiency that was 50-fold higher than that of the 2,200-bp control. In contrast, the 22-bp control sequence did not generate any gentamicin-resistant recombinants, indicating that the number of integration events was lower than the detection limit of 1 CFU per ~2 × 10^9 cells. In the ΔrecA strain, no transformants were detected for the 2,200-bp control, which relies on homologous recombination for integration, while the att site-containing plasmid integrated at a level similar to that in the wild-type strain (Fig. 2B). In both the wild-type and ΔrecA strains, integration of the att-containing plasmid at the chromosomal dif site was confirmed by PCR amplification of the dif region and Sanger sequencing (data not shown).

These data indicate that, similar to other dif sites, site-specific recombination at the L. pneumophila dif site is efficient and RecA independent.
Legionella Xer/dif is a single-recombinase system. We next aimed to identify the recombinase or recombinases that catalyze recombination at the Legionella dif site. No orthologs of XerC, XerD, XerS, or XerH have been identified in Legionella using sequence homology (29). However, HHpred analysis (36) revealed that the proteins encoded by six
L. pneumophila genes present in strain Lp02 (lpg0980, lpg0981, lpg1070, lpg1085, lpg1867, and lpg2057) show structural similarity to tyrosine-type recombinases (probability of >99%). Importantly, dif and xer deletion mutants in other species exhibit slight growth defects and filamentation but are viable (11, 13, 14, 18, 37), allowing for analysis of recombination in these deletion strains. We therefore repeated our intermolecular recombination assay in strains containing individual deletions of these six genes, along with a strain lacking the dif site. As expected, deletion of the dif site resulted in no detectable integration of the att site-containing plasmid. Deletion of lpg1867 also specifically abolished integration, while deletion of each of the other five recombinases had no significant effect (Fig. 3A).

We also performed rescue experiments using expression plasmids to confirm that the loss of recombination at the dif site in the Δlpg1867 strain was due to loss of Lpg1867 protein function and not any secondary mutations acquired during strain generation. Expression of Lpg1867 in the Δlpg1867 strain resulted in near-wild-type levels of recombination (Fig. 3B). In contrast, no recombination was detected with an empty vector control or when a point mutant (Y387F) of Lpg1867 (Lpg1867-Y387F) was expressed. The Y387F mutant is predicted to be catalytically inactive based on the equivalent catalytic tyrosine mutation in other tyrosine recombinases (10, 38, 39).

We next asked whether one of the other five putative tyrosine recombinases might be involved along with Lpg1867 in recombination at the dif site, despite having no effect upon individual deletion. To address this, we generated two multiple-deletion

---

**FIG 2** Site-specific recombination at the dif site is highly efficient and does not require RecA. (A) Schematic representation of the intermolecular recombination assay used to quantify att site-specific recombination in L. pneumophila. A 5.5-kb nonreplicating plasmid modified to include the 22-bp LME-1 att site (or control sequences) is transformed into L. pneumophila strain Lp02. If recombination between the plasmid att site and the chromosomal dif site takes place, the whole plasmid is integrated into the genome and gentamicin resistance is conferred to the cell. Gentamicin-resistant colonies are counted to quantify recombination. (B) Transformation efficiency (relative to that of a replicative plasmid control) of a nonreplicative plasmid containing the 22-bp att site, a 22-bp control sequence, or a 2,200-bp intergenic region of the Lp02 genome, transformed into the wild-type or ΔrecA Lp02 strains. Each dot shows the value obtained from a single experiment, and each horizontal line represents the geometric mean of 3 independent experiments.
strains: one with five putative recombinases deleted (all except lpg1867) and a pan-deletion strain with all six deleted. In our intermolecular recombination assay, integration of the att-containing plasmid was similar to wild-type levels in the 5-deletion strain and nonexistent in the pan-deletion strain (Fig. 4A). Overexpression of Lpg1867, but not Lpg1867-Y387F, again rescued the recombination defect (Fig. 4B).

As chromosome dimer resolution involves recombination between two dif sites on the same molecule, we also wanted to test whether Lpg1867 was sufficient to catalyze intramolecular recombination between two dif sites. To this end, we developed an intramolecular recombination assay by inserting a dif site-flanked kanamycin resistance (Kan') cassette at the normal dif locus of the Lp02 pan-deletion strain, such that recombination between the two flanking dif sites would result in loss of kanamycin resistance. To quantify excision of the reporter, we determined the proportion of kanamycin-resistant cells in the population after transforming the cells with wild-type or mutant Lpg1867. Overexpression of wild-type Lpg1867 in this strain resulted in a high level of cassette excision, while overexpression of the catalytic mutant resulted in an excision level similar to background level (Fig. 4C). Together, these data suggested that Lpg1867 does not require a partner recombinase for recombination at the dif site. However, to rule out the possibility that another unidentified recombinase or cofactor in L. pneumophila is also required, we performed a similar intramolecular excision assay in yeast. In this case, a dif site-flanked KanMX cassette was inserted at the HO locus of the Saccharomyces cerevisiae BY4741 strain, and recombination between the two dif sites was assessed by quantifying loss of Geneticin resistance (Fig. 4C). Overexpression of Lpg1867, but not Lpg1867-Y387F, again led to high levels of excision. Collectively, these data indicate that Lpg1867 is both necessary and sufficient to catalyze site-specific recombination at the L. pneumophila dif site.

Deletion of dif or lpg1867 induces filamentation and inhibits extracellular and intracellular growth. Loss of a functional chromosome dimer resolution system causes defects in cell division, which can result in a variety of phenotypes, including slow growth and filamentation. (The dif site was named for its deletion-induced filamentation phenotype.) Consistent with reported findings for dif or xer deletions in other species (11, 13, 14, 18, 37, 40), we observed that loss of dif or Lpg1867 in L. pneumophila led to filamentation in a subpopulation of cells (Fig. 5A). Notably, and consistent with a reduction in chromosome dimer formation in the absence of RecA (2, 41), we observed a decrease in filament formation upon deletion of recA in the Δlpg1867 background (Fig. 5A).
Deletion of \textit{dif} or \textit{lpg1867} also results in a visible reduction in colony size on solid media (Fig. 5B). Again, this defect is alleviated upon additional deletion of \textit{recA} in the \textit{Dlpg1867} background (Fig. 5B), consistent with the growth inhibition being caused by an inability to resolve chromosome dimers. The slow-growth phenotype is also present in the recombinase pan-deletion strain, but not the 5-deletion strain, and the defect is rescued by expression of wild-type \textit{Lpg1867}, but not the catalytic mutant (Fig. 5C), mirroring the effects of these deletions on recombination (Fig. 4A and B). Importantly, the \textit{dif} and \textit{lpg1867} deletions also result in diminished growth in human U937 macrophages (Fig. 5D), indicating that growth within host cells is also impacted by disruptions to \textit{dif}/Xer activity. Taken together, these data support a model in which a single site-specific recombinase, \textit{Lpg1867} (which we designate here as XerL), catalyzes recombination between two \textit{dif} sites to resolve chromosome dimers in \textit{Legionella}.

**LME-1 stability is modulated by the formation of a modified downstream \textit{dif} site.** We have previously shown that LME-1 integrates into the \textit{L. pneumophila} chromosome at what we now know to be the \textit{dif} site (33). LME-1 is stably integrated into the genome of strain Murcia-4983 (only excised in \(\sim\)1% of the population), despite being flanked by two \textit{att} sites (33). The sequence of the LME-1 \textit{att} site is identical to that of
the L. pneumophila dif site, but it encompasses only 22 bp of it. Recombination between this 22-bp sequence and the chromosomal dif site results in att site duplication, which generates an intact dif site upstream and a modified dif site downstream of LME-1 that contains three substitutions (Fig. 6A). We next asked whether these modifications to the downstream dif site contribute to the stability of the integrated form of LME-1. We modified our excision assay to use the kanamycin resistance cassette as a proxy for LME-1 and determined the effect of adding one or more of the LME-1 att

| Strain                  | Filament frequency (filaments/cells counted) |
|-------------------------|---------------------------------------------|
| wildtype                | 1.5% (49/3194)                              |
| Δdif                    | 11.1% (304/2741)                            |
| Δlpg1867                | 12.5% (408/3261)                            |
| Δlpg1867ΔrecA           | 3.9% (113/2897)                             |

FIG 5  dif and lpg1867 deletion strains exhibit filamentation and altered growth on solid media and during infection of host cells. (A) DAPI-stained cells of the indicated L. pneumophila strains showing filamentation upon deletion of the dif site or lpg1867 and rescue when recA is deleted in the Δlpg1867 background. The scale bars represent 10 μm. Frequency of filamentation was quantified using the MicrobeJ plugin for ImageJ (58) and is shown in the right panel. (B, left panel) CYET spread plates after 4 days of growth at 37°C with visible differences in colony size. (Right panel) Each colony radius is plotted as a separate data point, with the geometric mean shown by the purple line. (C) Distribution of colony sizes in the 5-deletion strain and pan-deletion strain (left panel) and in the pan-deletion strain expressing wild-type or mutant Lpg1867 (right panel). Data are plotted as in panel B. (D) Intracellular growth of wild-type and deletion strains in U937 macrophages. The data are plotted as the fold increase in CFU recovered after lysis from host cells at 2 h and 48 h postinfection relative to wild-type Lp02. Data show the average of three replicates for each strain from a representative experiment. Error bars represent SEM.
substitutions to the downstream dif site (Fig. 6B). We found that adding all three substitutions to the downstream dif site (to mimic integrated LME-1) resulted in markedly reduced excision compared to the reporter flanked by two wild-type dif sites, while one or two substitutions resulted in an intermediate level of excision (Fig. 6B). These data suggest that the short 22-bp att site of LME-1 likely contributes to the stability of its integrated form by preventing it from being flanked by two intact dif sites.

**Widespread distribution of xerL orthologs in *Legionella* and *Coxiella*.** Our discovery of the *Legionella* dif/Xer pathway, which had been missed in homology-based searches due to divergence from known CDR components, led us to next examine other bacteria in which no dif/Xer system has been identified to date. One possibility is that several species have XerL-like systems that were missed for the same reasons that they were missed in *Legionella*. To explore this possibility, we performed blastp and tblastn analyses using the *L. pneumophila* XerL protein as a query. These analyses identified several additional XerL orthologs across the order Legionellales, which includes *Legionella* sp. as well as the select agent *Coxiella burnetii* (Fig. 7). Phylogenetic analysis of the XerL-like recombinases indicates widespread distribution across Legionellales (Fig. S1). The amount of sequence divergence between XerL and the rest of the Xer protein family likely explains how this machinery has been missed by prior analyses.

The *Coxiella burnetii* xerL ortholog is encoded on a virulence plasmid. Despite the apparent vertical inheritance of xerL given its restriction within the order Legionellales, the *Coxiella burnetii* orthologs of xerL (here xerL_{cb}) reside not on the circular chromosome, but on a large ubiquitous plasmid that is critical for virulence (42–45). While Xer/dif systems have established roles in resolving aberrant plasmid dimers, these are the only Xer sequences within the *C. burnetii* genome. This raises the question as to whether this plasmid-based XerL functions to resolve plasmid dimers, chromosome dimers, or both.

To search for putative dif sites within the *Coxiella* genome, we performed a sequence similarity (blastn) search using the 29-nucleotide (nt) conserved sequence containing the *L. pneumophila* dif site. We identified a short sequence of 30 nt on the virulence plasmid immediately adjacent to XerL_{cb} (Fig. 8A). Using this second sequence as a query, we identified an additional site on the chromosome (Fig. 8A). This chromosomal sequence diverges from the plasmid site yet has several features consistent with a functioning dif
site. Like the *L. pneumophila* dif sequence, it is palindromic, highly conserved across isolates, and has no apparent coding potential.

To further investigate the potential role of this chromosomal site as a functional dif site, we examined its location relative to the point of inflection of GC skew (replication terminus). As we observed for dif(Lp), the *Coxiella burnetii* chromosomal site is always located opposite the predicted origin of replication and close to the cumulative GC skew maximum (Fig. 8B and Table 52). Taken together, this sequence not only shares homology to dif(Lp), but also maintains several hallmarks of a dif site involved in chromosome dimer resolution. Here, we will refer to it as dif(Cb).

The *Coxiella* XerL/dif system functionally complements a ΔxerL/dif mutant in *Legionella*. Having identified putative dif sites and xerL orthologs in *Coxiella burnetii*, we next wanted to assay their potential for chromosome dimer resolution. Considering that colony size correlates with dif/Xer function in *L. pneumophila* (Fig. 5), we generated a strain in which the endogenous dif site and xerL gene were replaced by the *C. burnetii* dif site and xerL and assessed growth on solid media. If the *C. burnetii* dif/XerL strains were
FIG 8 The *C. burnetii* *dif* site and *XerL* can functionally replace endogenous *dif*/*Xer* in *L. pneumophila*. (A) Sequence alignment of the *L. pneumophila* *dif* site with the plasmid and chromosomal *dif* sites of *Coxiella burnetii*. The gray arrows indicate the (Continued on next page)
not able to resolve dimers in their surrogate chromosome, then we would expect small colony sizes similar to those of the Δdif or ΔxerL strains. However, the colony size distribution of the strain containing XerL_Cb and the C. burnetii chromosomal dif site (dif_CbChr) was almost identical to that of wild-type L. pneumophila (Fig. 8C), while a strain encoding the predicted catalytic mutant XerL-Y389F displayed the small colony sizes associated with nonfunctional dif/Xer (Fig. 8C). This suggests that despite being encoded on the plasmid, XerL_Cb may catalyze recombination at dif_CbChr to resolve dimers of the Coxiiella chromosome. Interestingly, despite containing several substitutions (Fig. 8A), the Coxiiella plasmid dif site (dif_CbPl) in concert with XerL_Cb is also able to complement the small colony phenotype (Fig. 8C). Similarly, in our intramolecular recombination assay, the reporter cassette was excised when flanked by either dif_CbChr or dif_CbPl, although excision for the cassette flanked by dif_CbPl required overnight induction of XerL_Cb (Fig. 8D).

Together, these findings suggest that XerL_Cb may resolve both plasmid and chromosome dimers in Coxiiella, with the resolution of chromosome dimers uniquely dependent on plasmid-encoded XerL. XerL is encoded by all four virulence plasmids described in C. burnetii, and no additional Xer orthologs can be found on the chromosome. This suggests that chromosome dimer resolution and virulence plasmid maintenance are genetically linked in C. burnetii, with plasmid loss leading to loss of the chromosome dimer resolution pathway. Consistent with this model, we note that previously described instances of plasmidless strains—in which a subset of plasmid-like sequences are instead integrated onto the chromosome (46–48)—contain XerLCb within the integrated sequence. Our findings suggest that such integrants are likely selected for based on their ability to escape the fitness costs that a plasmidless strain would otherwise incur. Despite the sequence similarity between the dif sites located on the plasmid and chromosome and their shared dependence on XerL, the integration of plasmid-like sequences in the plasmidless strains is not at the chromosomal dif site. In fact, the plasmid fragments are located ~1 Mb from the terminus region, indicating that acquisition of the plasmid-like sequences did not originate with an IMEX-like plasmid integration event. To see if this phenomenon would recapitulate in our intermolecular recombination assay, we measured integration of a nonreplicative plasmid containing dif_CbPl into the genome of a strain containing dif_CbChr and found it to be almost 300-fold lower than when the genome contained dif_CbPl (Fig. 8E). Surprisingly, the efficiency of integration into the dif_CbChr-containing genome was also very low when the nonreplicative plasmid contained the identical dif_CbChr (Fig. 8E). One possible explanation is that highly efficient recombination between the two dif sites leads to rapid reexcision after integration, resulting in no visible gentamicin-resistant colonies. However, when we used a version of dif_CbChr with a shortened region of dyad symmetry, which is known to increase stability of integration for the LME-1 dif mimic (Fig. 6), transformation efficiency remained very low (Fig. 8E). These findings suggest that either shortened dyad symmetry does not stabilize dif_CbChr or the chromosomal dif site itself might be recalcitrant.
to invasion. The latter notion is consistent with the fact that no sequenced *Coxiella burnetii* genomes contain virulence plasmids as chromosomal integrants at the *dif* site.

**DISCUSSION**

The discovery of the *L. pneumophila* CDR system has solved two mysteries of the pathogen’s biology: the function of the conserved DNA sequence that is hijacked by LME-1 and how *Legionella* species are able to resolve chromosome dimers despite having no homologs to known *dif* sites and Xer recombinases. Similarly, no *dif*/Xer components had been identified previously in members of the other *Legionellales* family, *Coxiellaceae* (29), which includes the human pathogen *Coxiella burnetii*. Our discovery of *dif*/XerL in these species has uncovered an overlooked, highly important pathway that is conserved across an entire order. Other extensively studied dimensions of *Legionella* biology (e.g., translocated effectors) are not strong candidates for drug targets due to their redundancy. In contrast, we show that disruption of one protein in this pathway (XerL) is sufficient to dramatically restrict the growth of the pathogen.

With no apparent *dif* or Xer homologs in *Legionella*, the CDR system was instead unearthed through our investigation of LME-1 and its attachment site. In turn, the discovery of the *Legionella* *dif* site and XerL has allowed us to explore new LME-1 biology. Our findings are consistent with LME-1 belonging to the IMEX class of integrative mobile elements, which exploit the host CDR machinery for integration and in some cases excision (5, 28). The mechanistic details of LME-1 integration and excision and how they relate to the strategies of other IMEXs are interesting topics of future study. One intriguing aspect of LME-1 is the stability of its integrated form. Our reporter excision assay showed that the modified *dif* site and wild-type *dif* site that flank the integrated LME-1 recombine at a very low level and that the stability of integration is influenced by the three nucleotide substitutions present in the modified form. Another stably integrated IMEX, the gonococcal genomic island (GGI) of *Neisseria gonorrhoeae*, generates a similarly modified *dif* site (*dif*\textsubscript{gg}) at one end of its integrated form, while the wild-type *dif* (*dif*\textsubscript{wt}) remains intact at the other (23). As for many other species, CDR in *Neisseria* is facilitated by XerC and XerD homologs and is dependent on the DNA translocase FtsK, which pauses at the *dif*/Xer complex and activates XerD. Interestingly, FtsK does not stop at the *dif*\textsubscript{gg}/XerC/D complex, but rather translocates through it, which likely leads to disassembly of the complex and precludes recombination between the *dif*\textsubscript{gg} and *dif*\textsubscript{gg} (23). It is possible that this FtsK-dependent process contributes to the stability of the integrated form of LME-1, although we have yet to determine whether *Legionella* CDR requires its FtsK homolog (lpg1766). Our future investigative priorities include determining the involvement of FtsK in *L. pneumophila* CDR and its contribution to the stability of LME-1. It is possible that even if *Legionella* CDR is FtsK dependent, LME-1 might circumvent FtsK for integration and excision, as is the case for several other IMEX elements (28). One IMEX of *Vibrio cholerae* escapes FtsK dependence by encoding its own XerD-activating factor, XafT (49). It will be interesting to see whether LME-1 employs a similar strategy and encodes a factor that can activate XerL.

The *Legionella* and *Coxiella* *dif*/Xer systems are distinct in several ways from those characterized previously in other bacterial species. The *dif* site arms consist of long inverted repeats (10 and 12 bp, respectively) with no interruptions, whereas other *dif* sites show only partial dyad symmetry between the two arms, even for single-recombinase systems. Additionally, in *L. pneumophila*, *dif* and xerL are separated by >400 kb in the genome, whereas in the other bacterial single-recombinase systems (i.e., *dif*\textsubscript{x}/Xer\textsubscript{S} of streptococci and lactococci and *dif*\textsubscript{H}/Xer\textsubscript{H} of *Campylobacter* and *Helicobacter*), the *dif* site is always near the xer gene, indicating they may have been acquired as a single module (3, 16, 17, 29). The location of the *C. burnetii* xerL ortholog on its virulence plasmid is also unusual. While many plasmids use Xer machinery for their own dimer resolution, this is the first instance we are aware of where a plasmid-based Xer recombinase is necessary for chromosome dimer resolution.

Given the general importance of chromosome dimer resolution to bacterial fitness
and the extent to which intracellular growth is compromised in \textit{L. pneumophila} \textit{Δxerl} or \textit{dif} mutants, we anticipate that plasmid loss in \textit{Coxiella} incurs a significant fitness cost through the loss of the \textit{xerl} recombinase. Consistent with this model, isolates that have lost their plasmid appear to maintain \textit{dif}/\textit{XerL} activity by including \textit{xerl} within the fragments of plasmid sequence found to be integrated on the chromosome. One interpretation of these results is that integrants that do not maintain \textit{xerl} chromosomally are rapidly selected against and not recovered. In light of our data, it will be important to distinguish which phenotypes previously associated with directed plasmid loss (44) reflect virulence-specific defects (associated with loss of specific virulence factors, such as effectors) and which phenotypes result from the loss of \textit{dif}/\textit{XerL} machinery.

**MATERIALS AND METHODS**

**Strains and plasmids.** The \textit{L. pneumophila} Lp02 strain was used as the background for all \textit{L. pneumophila} strains generated in this study. The \textit{L. pneumophila} Murcia-4983 strain, which harbors LME-1, is an environmental isolate collected during the 2001 outbreak in Murcia, Spain (50). All \textit{L. pneumophila} strains were grown at 37°C in N-(2-acetamido)-2-aminoethanesulfonic acid (ACES) buffered yeast extract supplemented with 100 μg/mL thymidine (AYET) and on charcoal AYET (CYET) agar plates. Deletion and mutant strains were made using a scar-free suicide cassette method described recently (51) with some minor modifications. This protocol uses a 2-stage approach to introduce mutations. First, linear DNA containing a \textit{mazF-Kan}’ cassette flanked by ~1.5 kb chromosomal homology arms is introduced into the target plasmid via transformation. The \textit{mazF-Kan}’ cassette provides both positive (Kan') and negative (IPPG [isopropyl-thiogalactopyranoside]-inducible \textit{mazF} toxin) selection markers. Cells in which the \textit{mazF} cassette has been integrated via homologous recombination are selected by plating on medium containing kanamycin. The resulting colonies are screened on kanamycin- and IPTG-containing plates to identify cells that have \textit{mazF-Kan}’ incorporated (growth on kanamycin, no growth on IPTG). In the second stage, these integrant strains are transformed with linear DNA containing only the homology arms with the desired mutation. Homologous recombination leads to excision of the cassette, and in cells in which the cassette was lost are selected on plates containing IPTG. Colonies are again screened on kanamycin- and IPTG-containing plates to identify those in which the cassette was removed (i.e., growth on IPTG, no growth on kanamycin). All mutations made in this study were verified by PCR screening and Sanger sequencing.

The reporter strains used for the intramolecular recombination (reporter excision) assays were generated by introducing linear DNA containing the \textit{dif-Kanr}-\textit{dif} cassette (flanked by ~1.5-kb homology arms around the chromosomal \textit{dif} site) into the Lp02 \textit{Argp1867} strain or pan-deletion strain. Strains containing the cassette were selected on plates containing kanamycin, and the correct integration of the cassettes was verified by Sanger sequencing. Similarly, for the yeast-based excision assay, linear DNA containing the \textit{dif-KanMX-dif} cassette (flanked by homology arms around the HO locus) was transformed into \textit{S. cerevisiae} strain BY4741 using the high-efficiency polyethylene glycol (PEG)-LiAc method (52). Integrants were selected on YPD agar plates containing Geneticin and verified by Sanger sequencing.

The plBJ4648-based plasmids used for the intermolecular recombination assays were constructed by cloning the 22-bp \textit{att} site sequence, a control 22-bp sequence, or 2,200 bp of intergenic \textit{Lp02} sequence into \textit{pBJ4648} digested with XhoI and ApaI. The gentamicin-resistant plasmid pJB1806-GmR was generated by first amplifying the gentamicin resistance (Gmr) gene and promoter from \textit{pJB4648} and cloning it in \textit{pJB4648} digested with XhoI and ApaI. The chromosomally integrated \textit{pJB1806-GmR} was generated by transforming either carrier strain with pBJ4648 and cloning it into the chromosome. The \textit{dif} cassette was introduced into the \textit{L. pneumophila} strain via natural transformation. The \textit{pJB1806-GmR} plasmid was linearized by digestion with BgIII, and the 

**GC skew analysis.** The cumulative GC skew maximum value for each \textit{Legionella} genome (see Table S1 in the supplemental material) or \textit{C. burnetii} genome (Table S2) was determined using the GenSkew java app (https://genskew.cs.univie.ac.at) with default settings. The whole-genome GC skew profiles and \textit{dif} site locations for seven \textit{L. pneumophila} isolates, seven \textit{Legionella} species, and seven \textit{C. burnetii} strains, selected to show a range of GC skew maximum locations relative to the origin, were plotted using Circos (55). The \textit{GC skew} profiles of these genomes were normalized for genome length. The selected \textit{L. pneumophila} isolates and \textit{Legionella} species (with GenBank accession numbers in parentheses) are \textit{L. pneumophila} Philadelphia-2 \textit{(CP015929.1)}, Philadelphia-1 \textit{(AE017354.1)}, E7-O \textit{(CP015954.1)}, Paris \textit{(CP0628336.1)}, NY23-D7705 \textit{(CP021261.1)}, NCTC11985 \textit{(LT906452.1)}, and \textit{L. pneumophila} subsp. \textit{fraseri} strain D-4058 \textit{(CP021277)}. and \textit{Legionella} species \textit{L. copenhagenii} L18-01051 \textit{(CP041668.1)}, \textit{L. micdadei} \textit{(LN614830.1)}, \textit{L. spiretens} \textit{(NCTC11990 LT906457.1)}, \textit{L. saintheleniae} \textit{(NCTC12450 LR134178.1)}, \textit{L. longbeachae} \textit{(N650140.1)}, \textit{L. oakridgensis} \textit{(NCTC11531 LR134286.1)} and \textit{L. lompovii} \textit{(LN614827.1)}. The selected \textit{C. burnetii} isolates are \textit{LN614828.1} \textit{(CP009173.1)}, \textit{RSA 493} \textit{(AE016628.3)}, \textit{CbuG, Q212} \textit{(CP001019.1)}, \textit{Schelperling} \textit{(CP014563.1)} and \textit{CbuK, Q154} \textit{(CP001020.1)}.

**Recombination assays.** For the intermolecular recombination (plasmid integration) assays, each strain was transformed with a \textit{pBJ4648}-based plasmid containing the indicated \textit{att}, \textit{dif}, or control sequences as indicated. The strains were also transformed with the replicative plasmid pJB1806-GmR to

\[ \text{http://genskew.cs.univie.ac.at} \]
control for strain-to-strain differences in general transformation efficiency. Plasmid concentrations were determined using a Quant-IT PicoGreen double-stranded DNA (dsDNA) kit (Invitrogen), and 200 ng of each plasmid was electroporated into 2 optical density units (ODU) of L. pneumophila cells as described previously (56). After an 8-h recovery period in liquid AYET medium shaking at 37°C, dilutions of each sample were plated onto CYET plates supplemented with 15 μg/mL gentamicin and incubated at 37°C for 4 days. Colonies were counted using GeneTools analysis software (Syngene), and the results were plotted using GraphPad Prism version 9.2.0.

For the intramolecular recombination (excision) assay in Legionella, strains containing the kanamycin resistance cassette were transformed with tet-on-pJB1806-GmR containing wild-type Lpg1867, Lpg1867-Y387F, XerLrc, XerLrc-Y389F, or empty vector by electroporation as described above. Transformants were selected on CYET plates containing both gentamicin and kanamycin. The transformant colonies were then collected in AYE and spread onto nonselective plates and plates containing kanamycin to quantify the proportion of the population that had lost kanamycin resistance. For the intramolecular recombination (excision) assay in yeast, S. cerevisiae strain BY4741 containing the dif-KanMX-dif cassette was transformed with yeast vector pAG413GPD (57) expressing wild-type or mutant Lpg1867 or empty vector. The percentage of excision was quantified by determining the proportion of the transformants that had lost Geneticin resistance by plating on selective and nonselective plates.

**Fluorescence microscopy.** Overnight cultures of the Lp02, Lp02 Δdif, Lp02 Δlpg1867, and Lp02 Δlpg1867 ΔrecA strains were grown from a 2-day-old patch in AYET medium. Bacteria were collected at postexponential phase (OD at 600 nm [OD_{600}] of 4.0 to 4.5). Approximately \( \sim 1 \times 10^7 \) cells (1 ODU) from each strain were washed once with 1 mL of 1 x phosphate-buffered saline (PBS) before a final resuspension in 500 μL of 1 x PBS. The bacteria were stained with 15 μg/mL of DAPI (4',6-diamidino-2-phenylindole) (Roche) for 30 min at room temperature and imaged using a 63 x oil immersion lens on a Zeiss Axio Imager.M2 microscope. The frequency of filamentation was quantified using the Microbel plugin for ImageJ (58).

**Colony size assay.** The distribution of colony size was determined by resuspending 2-day-old patches of the indicated strains in AYET. The OD_{600} of each cell suspension was determined, and the appropriate dilutions made so that \( \sim 200 \) to 300 cells were added to each plate. Spread plates were incubated at 37°C for 4 days before imaging using a SynGene system with GeneSnap software (Syngene). Colony diameter was determined using ColTapp automated image analysis software (59) and plotted using GraphPad Prism version 9.2.0.

**Bacterial infection of U937 macrophages.** Tetradecanoyl phorbol acetate (TPA)-differentiated U937 cells were seeded in a 24-well plate at \( 4 \times 10^5 \) cells per well in 500 μL of RPMI supplemented with glutamine and 10% heat-inactivated fetal bovine serum. Cells were incubated at 37°C with 5% CO₂ for a further 2 days. At 2 h and 48 h postinfection, cells were lysed with 0.02% saponin and then plated on CYET to determine the number of bacterial CFU.

**Bioinformatic analysis of XerL.** The amino acid sequence of XerL from L. pneumophila was used to uncover putative orthologs of XerL. XerL was queried against the NCBI Legionellales database (taxid 118969) using both blastp and tblastn with the default settings (60), in addition to a tblastn search of additional Legionella species that were deposited in the NCBI Sequence Read Archive (SRA) (61, 62). The resulting putative XerL orthologs can be found in Table S3. An unrooted phylogeny of Xer proteins was generated using the amino acid sequences from representative XerL orthologs, in addition to reference Xer sequences (Table S3). The amino acid sequences were aligned using MUSCLE (63), and the tree was generated with the FastTree (v.2.1.11) plugin on Geneious Prime using the default settings (64). The resulting tree was visualized using the Interactive Tree of Life (iTOL) server (v.6) (65).

A rooted phylogeny of all putative XerL orthologs was generated as described above, using XerD from Escherichia coli as the root.

**Data availability.** All strains and plasmids are available upon request. Sequences and locations of dif sites are listed in figures and Table S1 and Table S2. The accession numbers and amino acid sequences for Xer proteins are listed in Table S3.

**Supplemental Material**

Supplemental material is available online only.

**FIG S1**, EPS file, 0.1 MB.

**TABLE S1**, XLSX file, 0.03 MB.

**TABLE S2**, XLSX file, 0.01 MB.

**TABLE S3**, XLSX file, 0.03 MB.

**Acknowledgments**

We thank members of the Ensminger laboratory for suggestions and careful reading of the manuscript and Jordan Lin and John MacPherson for help with the recombination assays and HHPrep analysis. We also thank Hayley Newton for the kind gift of genomic DNA from C. burnetii strain Nine Mile phase II.
REFERENCES

1. Barre FX, Sabbale B, Michel B, Aroyo M, Robertson M, Sherratt D. 2001. Circles: the replication-recombination-chromosome segregation connection. Proc Natl Acad Sci U S A 98:8189–8195. https://doi.org/10.1073/pnas.110008998.

2. Steiner WW, Kueppler PL. 1998. Sister chromatid exchange frequencies in Escherichia coli analyzed by recombination at the dif resolvasse site. J Bacteriol 180:6269–6275. https://doi.org/10.1128/JB.180.23.6269-6275.1998.

3. Castillo F, Benmouamed A, Sazmati G. 2017. Xer site specific recombination: double and single recombination systems. Front Microbiol 8:453. https://doi.org/10.3389/fmicb.2017.00453.

4. Diagne CT, Salhi M, Crozat E, Salomé L, Cornet F, Rousseau P, Tardin C. 2014. TPM analyses reveal that FtsK contributes both to the assembly and the activation of the XerC-dif recombination synapse. Nucleic Acids Res 42:1721–1732. https://doi.org/10.1093/nar/gkt1024.

5. Midonet C, Barre F-X. 2014. Xer site-specific recombination: promoting vertical and horizontal transmission of genetic information. Microbiol Spectr 2:MDNA3-0056-2014. https://doi.org/10.1128/microbiolspec.MDNA3-0056-2014.

6. Thanbichler M. 2010. Synchroization of chromosome dynamics and cell division in bacteria. Cold Spring Harb Perspect Biol 2:a000331. https://doi.org/10.1101/cshperspect.a000331.

7. Austin S, Ziese M, Sternberg N. 1981. A novel role for site-specific recombination in maintenance of bacterial replicons. Cell 25:729–736. https://doi.org/10.1016/0092-8674(81)90180-x.

8. Summers DK, Sherratt DJ. 1984. Multimerization of high copy number plasmids causes instability. CoIE1 encodes a determinant essential for plasmid monomerization and stability. Cell 36:1097.0092-8674(84)90060-6

9. Colomos SD, Sykora P, Sazmati G, Sherratt DJ. 1990. Recombination at CoIE1 cer requires the Escherichia coli xerC gene product, a member of the lambda integrase family of site-specific recombinases. J Bacteriol 172:6973–6980. https://doi.org/10.1128/jb.172.12.6973-6980.1990.

10. Blakely G, May G, McCulloch R, Arciszewska LK, Burke M, Lovett ST, Sherratt DJ. 1993. Two related recombinases are required for site-specific recombination at dif and cer in the e. coli K12. Cell 75:351–361. https://doi.org/10.1016/0092-8674(93)80076-Q.

11. Blakely G, Colomos S, May G, Burke M, Sherratt D. 1991. Escherichia coli XerC recombinase is required for chromosomal segregation at cell division. New Biol 3:789–798.

12. Clerget M. 1991. Site-specific recombination promoted by a short DNA segment of plasmid R1 and by a homologous segment in the terminus chromosome of Escherichia coli. New Biol 3:780–788.

13. Kueppel PL, Henson JM, Dieck L, Tecklenburg M, Dircks L, Tecklenburg M, Lim DF. 1991. A recA-independent recombination site in the terminus region of the chromosome of Escherichia coli. New Biol 3:799–811.

14. Debowski AW, Carney C, Verbrughhe P, Nilsson H-O, Gauntlett JC, Fulurija V, Cornet F, Rousseau P, Tardin C. 2005. The single-stranded genome of phage CTX is the form used for integration into the genome of Vibrio cholerae. Mol Cell 19:559–566. https://doi.org/10.1016/j.molcel.2005.07.002.

15. Dan B, Martínez E, Midonet C, Barre F-X. 2013. Integrative mobile elements exploiting Xer recombination. Trends Microbiol 21:23–30. https://doi.org/10.1016/j.tim.2012.10.003.

16. Clerget M. 2001. Analysis of the terminus region of the Caulobacter crescentus chromosome and identification of the dif site. J Bacteriol 183:6016–6019. https://doi.org/10.1128/JB.00330-06.

17. Le Bourgeois P, Bugarel M, Campo N, Daveran-Mingot M-L, Labonté J, Forterre P, Serre M-C. 2010. Evidence for a Xer/dif system for chromosome recombination in Archaea. PLoS Genet 6:e1001166. https://doi.org/10.1371/journal.pgen.1001166.

18. Hassan F, Kamruzzaman M, Mekalanos JF, Faruque SM. 2010. Satellite phage TLPc{double}e enables toxigenic conversion by CTX phage through dif site alteration. Nature 467:982–985. https://doi.org/10.1038/nature09469.

19. Cortez D, Quevillon-Cheruel S, Gribaldo S, Desnoeus N, Sezovov G. 2010.Evidence for a Xer/dif system for chromosome recombination in Archaea. PLoS Genet 6:e1001166. https://doi.org/10.1371/journal.pgen.1001166.

20. Viguier C, Sadowski A, Gribaldo S, Desnoeus N, Sezovov G. 2010. Evidence for a Xer/dif system for chromosome recombination in Archaea. PLoS Genet 6:e1001166. https://doi.org/10.1371/journal.pgen.1001166.

21. Duggin IG, Dubarry N, Bell SD. 2011. Replication termination and chromo-some dimer resolution in the archaeon Sulfolobus solfataricus. EMBO J 30:145–153. https://doi.org/10.1038/emboj.2010.301.

22. Sciochetti SA, Piggot PJ, Blakely GW. 2001. Identi fication and characteriza- tion of the dif site from Bacteri subtilis. J Bacteriol 183:1058–1068. https://doi.org/10.1128/JB.183.3.1058-1068.2001.

23. Val MAE, Kennedy SP, El Karoui M, Bonne L, Chevalier F, Barre F-X. 2008. FtsK- dependent dimer resolution on multiple chromosomes in the pathogen Vibrio cholerae. PLoS Genet 4:e1000201. https://doi.org/10.1371/journal.pgen.1000201.

24. Forterre P, Serre M-C. 2010. Evidence for a Xer/dif system for chromosome recombination in Archaea. PLoS Genet 6:e1001166. https://doi.org/10.1371/journal.pgen.1001166.

25. Val M-E, Bischerour J, Barre F-X. 2011. VGphi integration and excision mecha-nisms contribute to the genetic diversity of Vibrio cholerae epidemic strains. Proc Natl Acad Sci U S A 108:2516–2521. https://doi.org/10.1073/pnas.1017061108.

26. Midonet C, Las, B, Paly E, Barre F-X. 2014. XerD-mediated FtsK-independent integration of TLPc{double}e into the Vibrio cholerae genome. Proc Natl Acad Sci U S A 111:16848–16853. https://doi.org/10.1073/pnas.140407111.

27. Fournes F, Crozat E, Pages C, Tardin C, Salomé L, Cornet F, Rousseau P. 2016. FtsK translation permits discrimination between an endogenous and an imported Xer/dif recombination complex. Proc Natl Acad Sci U S A 113:7882–7887. https://doi.org/10.1073/pnas.1523178113.

28. Qin W, Kamruzzaman M, Mekalanos JF, Faruque SM. 2010. Satellite phage TLPc{double}e enables toxigenic conversion by CTX phage through dif site alteration. Nature 467:982–985. https://doi.org/10.1038/nature09469.

29. Hendricks EC, Szerlong H, Hill T, Kuempel P. 2000. Cell division, guillotin ing of dimer chromosomes and SOS induction in resolution mutants (dif, xerC and xerD) of Escherichia coli. Mol Microbiol 30:1461–1473. https://doi.org/10.1046/j.1365-2958.2000.01920.x.
