Title

Control of Recombination Directionality by the Listeria Phage A118 Protein Gp44 and the Coiled-Coil Motif of Its Serine Integrase

Journal

Journal of Bacteriology, 199(11)

ISSN

0021-9193

Authors

Mandali, Sridhar
Gupta, Kushol
Dawson, Anthony R
et al.

Publication Date

2017-06-01

DOI

10.1128/jb.00019-17

Peer reviewed

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Bacteriology®

RESEARCH ARTICLE

Control of Recombination Directionality by the *Listeria* Phage A118 Protein Gp44 and the Coiled-Coil Motif of Its Serine Integrase

Sridhar Mandali,[a] Kushol Gupta,[b] Anthony R. Dawson,[a]* Gregory D. Van Duyne,[b] Reid C. Johnson[a,c]

Department of Biological Chemistry, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California, USA[a]; Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA[b]; Molecular Biology Institute, University of California, Los Angeles, Los Angeles, California, USA[c]

ABSTRACT The serine integrase of phage A118 catalyzes integrative recombination between *attP* on the phage and a specific *attB* locus on the chromosome of *Listeria monocytogenes*, but it is unable to promote excisive recombination between the hybrid *attL* and *attR* sites found on the integrated prophage without assistance by a recombination directionality factor (RDF). We have identified and characterized the phage-encoded RDF Gp44, which activates the A118 integrase for excision and inhibits integration. Gp44 binds to the C-terminal DNA binding domain of integrase, and we have localized the primary binding site to be within the mobile coiled-coil (CC) motif but distinct from the distal tip of the CC that is required for recombination. This interaction is sufficient to inhibit integration, but a second interaction involving the N-terminal end of Gp44 is also required to activate excision. We provide evidence that these two contacts modulate the trajectory of the CC motifs as they extend out from the integrase core in a manner dependent upon the identities of the four *att* sites. Our results support a model whereby Gp44 shapes the Int-bound complexes to control which *att* sites can synapse and recombine.

IMPORTANCE Serine integrases mediate directional recombination between bacteriophage and bacterial chromosomes. These highly regulated site-specific recombination reactions are integral to the life cycle of temperate phage and, in the case of *Listeria monocytogenes* lysogenized by A118 family phage, are an essential virulence determinant. Serine integrases are also utilized as tools for genetic engineering and synthetic biology because of their exquisite unidirectional control of the DNA exchange reaction. Here, we identify and characterize the recombination directionality factor (RDF) that activates excision and inhibits integration reactions by the phage A118 integrase. We provide evidence that the A118 RDF binds to and modulates the trajectory of the long coiled-coil motif that extends from the large carboxyl-terminal DNA binding domain and is postulated to control the early steps of recombination site synapsis.

KEYWORDS site-specific DNA recombination, serine integrase, recombination directionality factor, *Listeria monocytogenes*

Received 11 January 2017 **Accepted** 6 March 2017

Accepted manuscript posted online 13 March 2017

Citation Mandali S, Gupta K, Dawson AR, Van Duyne GD, Johnson RC. 2017. Control of recombination directionality by the *Listeria* phage A118 protein Gp44 and the coiled-coil motif of its serine integrase. J Bacteriol 199:e00019-17. https://doi.org/10.1128/JB.00019-17.

Editor Tina M. Henkin, Ohio State University

Address correspondence to Reid C. Johnson, rcjohnson@mednet.ucla.edu.

* Present address: Anthony R. Dawson, Department of Medical Microbiology & Immunology, University of Wisconsin, Madison, Madison, Wisconsin, USA.

Temperate phage genomes integrate into and excise out of bacterial chromosomes using phage-specific site-specific recombination systems. Many phage integrases are members of a subfamily of serine recombinases (SRs) that are known as large SRs due to their very long C-terminal DNA binding regions relative to other SRs (1–4). In the few serine integrase systems that have been studied, phage-encoded recombination

directionality factors (RDFs) that function together with their cognate integrases control the direction of the recombination reaction, integration or excision, and thus perform an essential function in the developmental pathways of lysogenic phage (5–9). The RDFs of serine integrases that have been studied thus far function mechanistically differently than the RDFs (or Xis) proteins of tyrosine integrases, such as those from phages λ, HP1, P2, P22, and L5, which are largely DNA architectural proteins (10–14).

Serine integrases are being used in a variety of genetic engineering and synthetic biology applications (15–22). A major attribute of serine integrases is that they perform unidirectional recombination, as regulated by their dedicated RDF. However, the mechanism underlying how the RDF controls directional recombination is poorly understood. It is anticipated that they will enjoy even greater utility in different applications as the regulation of these reactions becomes more fully understood at the molecular level.

We have been studying the phage A118 serine integrase from *Listeria monocytogenes* (23, 24). A118 integrates into the *comK* gene within the *L. monocytogenes* chromosome by recombination between the phage *attP* and chromosome *attB* loci to generate the hybrid sites *attL* and *attR* (25). This inactivates the *comK* gene, which in *Bacillus subtilis* is a transcriptional regulator of genes involved in DNA uptake (26). Recent work has shown that *comK* is an important regulator of essential virulence genes in *L. monocytogenes* (27). Prophage excision is induced during growth of *L. monocytogenes* within phagosomes of macrophages, and precise excision restoring functional *comK* is required for escape from the phagosome into the cytoplasm of the mammalian host.

Serine integrases bind as chemically inactive dimers to each partner *att* site, either *attB* and *attP* for integration or *attL* and *attR* for excision, and are remodeled into a tetramer during formation of the active synaptic complex (1, 28, 29). As demonstrated for other SRs, DNA exchange is believed to be mediated by double-strand DNA cleavages at the centers of each *att* site, translocation of one synaptic pair of subunits relative to the other within the tetramer, and religation of the four DNA strands (30–33). The chemical and protein isomerization events largely involve the N-terminal catalytic domain and long oligomerization helix in each subunit, which bear clear similarities to well-studied members of the small SR family. In contrast, the C-terminal DNA binding regions of serine integrases are much larger than in small SRs that mediate DNA inversion, deletion, or transposition and are believed to play a unique and critical role in dictating the directionality of the reaction (1, 2).

A crystal structure of the 320-amino-acid residue of C-terminal DNA binding region from the A118-related phage L1 in complex with an *attP* half site showed two separate subdomains bound over the 25-bp DNA segment (24). The more N-terminal recombinase domain (RD) binds over the *att* sequence that is proximal to the central crossover region, and the zinc ribbon domain (ZD) binds the distal end of the *att* site (illustrated in Fig. 6A). Comparisons of the 50-bp *attP* and 40-bp *attB* sequences led Rutherford et al. (24, 28) to propose that the position of the ZD on *attB* is shifted 5 bp closer to the RD, which can be accommodated by the flexible peptide chain connecting the two folds. The most striking feature of the structure was the long coiled-coil (CC) motifs that protruded from the ZD. The flexible CC motifs exhibited different trajectories as they extended out from the protein body in each of the four DNA complexes in the crystal asymmetric unit. Structural modeling suggested different architectures for the different *att* site complexes and the potential for the CC domains to both enhance and inhibit synapsis through *cis* and *trans* interactions. These models are consistent with earlier experimental data on directional synapsis by ϕC31 and Bxb1 serine integrases and the importance of amino acid residues within the CC motif region (24, 34–36). It was further proposed that binding of RDFs to serine integrases may alter the trajectory of the CC motifs, thereby regulating directional recombination (24, 28).

In this report, we identify and characterize the RDF for the phage A118 serine integrase. Like other RDFs for serine integrases that have been studied, the A118 RDF not only promotes the *attL* × *attR* excision reaction but inhibits the *attB* × *attP*

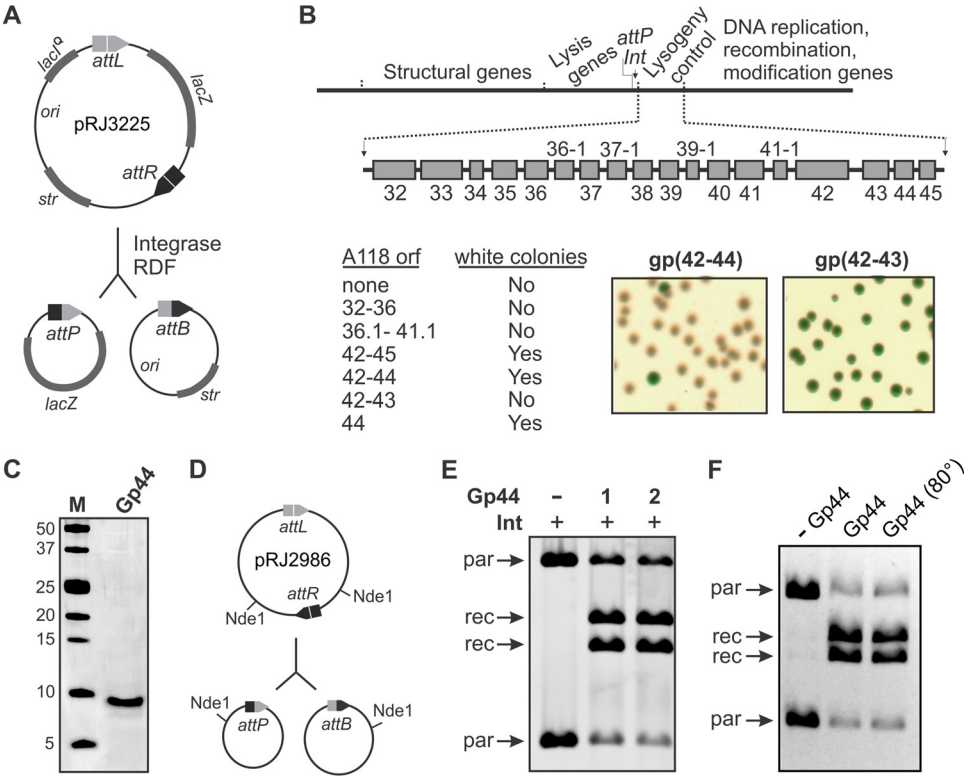

FIG 1 Identification of Gp44 as the A118 RDF. (A) Schematic of the *attL* × *attR* deletion reaction on the *in vivo* reporter pRJ3225. (B) Top, diagram of the phage A118 genome highlighting genes involved in the control of lysogeny. Integrase is encoded by *gp31*, and the RDF is encoded by gene *gp44*. Bottom, results of *in vivo* deletion assays testing various A118 regulatory genes. Formation of white colonies indicates that the *lacZ* gene has been lost due to *attL* × *attR* recombination forming a DNA circle lacking an origin of replication. On the right are representative images of colonies; the presence of mostly white colonies when cells were grown with A118 genes *gp42* to *gp44* indicates the presence of the RDF, whereas the absence of white colonies when cells were grown with genes *gp42* and *gp43* indicates no recombination. (C) Coomassie blue-stained SDS-PAGE gel of purified ^{His}Gp44 after chromatography on Ni-NTA, SP, and heparin. M, low-molecular-mass markers with sizes in kilodaltons noted. (D) Schematic of the *in vitro attL* × *attR* deletion reaction on pRJ2986. Products were digested with NdeI prior to electrophoresis; parental (par) and recombinant (rec) DNA fragments obtained after digestion are indicated in panels E and F. (E) *In vitro* excision reaction with purified Gp44 and Int. Reactions were performed without Gp44 (−), with ^{His}Gp44 (1), or with Gp44 purified from an Mxb intein fusion (2). (F) Gp44 is active after heat denaturation. *In vitro attL* × *attR* reactions were performed without Gp44 (−Gp44), with Gp44 (Gp44), and with Gp44 incubated at 80°C for 10 min [Gp44 (80°)].

integration reaction. We show that the A118 RDF binds to the integrase to generate electrophoretically distinct *att*-specific architectures. Inhibition of integration requires binding of the RDF to the CC motif on integrase. Activation of excision requires an additional interaction involving the N-terminal residues of the RDF. Our data support a model where the RDF modulates the trajectories of the CC motifs by binding to the CC and to a second region on integrase in an *att*-site specific manner.

RESULTS

Identification of Gp44 as the RDF for the A118 integrase. We asked whether an RDF function for phage A118 could be provided by a gene within the regulatory region between the integrase gene (*gp31*) and the replication, recombination, and modification genes on the phage genome (Fig. 1B). Segments of this 18-gene region (25) were cloned onto a multicopy plasmid and introduced into an *Escherichia coli* strain containing a *lacZ* reporter for *attL* × *attR* deletions (Fig. 1A) and expressing integrase (Int). RDF-dependent excision reactions were identified by white (Lac$^-$) colonies among the transformants on 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) plates. Segments containing *gp42* to *gp44* but not *gp42* to *gp43* enabled excision, pointing to *gp44* as the A118 gene encoding the RDF function (Fig. 1B). The presence of the *gp44* gene alone, as represented by pRJ3187 (Table 1), also stimulated excision.

TABLE 1 Plasmids and strains used in this study

Plasmid or strain	Description[a]	Reference or source
Plasmids		
pRJ2184	pET15b, A118 integrase (*gp31*) between NdeI and BamHI	23
pRJ2990	pET11a, A118 *gp44* with N-terminal 6His tag between NdeI and BamHI	This work
pRJ2823	pET15b, A118 integrase (residues 158–452) between NdeI and BamHI	23
pRJ2214	pBR322, *attP* (400 bp) into EcoRI	23
pRJ2215	pBR322, *attB* (200 bp) into SalI	23
pRJ2191	pBR322, *attL* (200 bp) between EcoRI and SalI	This work
pRJ2193	pBR322, *attR* (200 bp) between EcoRI and SalI	This work
pRJ2986	pBR322, *amp* DR(*attL-lacZ-attR*)	This work
pRJ3225	pMS421, *str* or *spc*, DR(*attL-lacZ-attR*)	This work
pRJ3244	pBR322, *amp* DR(*attB-lacZ-attP*)	This work
pRJ3245	pACYC184, *cat, lacPuv5-int* between SalI and XbaI	This work
pRJ3239	pET11a, *gp44*Δ(1–5) with N-terminal 6His tag between NdeI and BamHI	This work
pRJ3259	pRJ2184, *int*Δ(342–416)	This work
pRJ3002	pRJ2184, *int*Δ(381–383)	This work
pRJ3003	pRJ2184, *int*Δ(369–371)	This work
pRJ3187	pUC18, A118 *gp44* between PstI and SalI	This work
pRJ3188	pUC18, A118 *gp32–gp36* between PstI and SalI	This work
pRJ3189	pUC18, A118 *gp36.1–gp41.1* between PstI and SalI	This work
pRJ3190	pUC18, A118 *gp42–gp43* between PstI and SalI	This work
pRJ3191	pUC18, A118 *gp42–gp44* between PstI and SalI	This work
pRJ3192	pUC18, A118 *gp42–gp45* between PstI and SalI	This work
pGBKT7	Yeast two-hybrid DNA binding domain vector	Clontech
pGAD GH	Yeast two-hybrid activation domain vector	Clontech
pRJ3017	pGBKT7, A118 integrase between NdeI and BamHI	This work
pRJ3014	pGBKT7, A118 integrase (residues 158–452) between NdeI and BamHI	This work
pRJ3018	pGAD GH, A118 *gp44* between BamHI and XhoI	This work
pGBD-NPFSDx3	pGBD C1 (48), *STE2*-NPFSDx3	49
pGBD-NPASDx3	pGBD C1 (48), *STE2*-NPASDx3	49
pGAD-SLA1	pGAD (48), *SLA1*	49
pMS421	pSC101 origin, *str* or *spc*, *lacI*^{q}	M. Susskind
pGV2710	pETDuet (Novagen) with *gp44* + intein^{Mxe}-6His tag	This work
Bacterial and yeast strains		
DH5α	*E. coli* Δ*lacU169* ϕ80dΔ*lacZ*M15 *recA1 endA1 hsdR17 deoR glnV44 thi-1 gyrA96 relA1*	Laboratory collection
RJ3387	*E. coli* BL21(DE3) *endA8*::*tet fis*::*kan-767*	Laboratory collection
DP-L3689	*L. monocytogenes* (10403::A118)	R. Calendar
DP-L 4056	*L. monocytogenes* (10403S phage cured)	R. Calendar
PJ69-4A	*S. cerevisiae MAT***a** *trp1-901 leu2-3,112 ura3-52 his3-200* Δ*gal4* Δ*gal80 LYS2*::*GAL1-HIS3 GAL2-ADE2 met2*::*GAL7-lacZ*	48

[a]DR, *att* sites in direct repeat orientation.

The Gp44 protein with an N-terminal 6His tag was overexpressed in *E. coli* and purified by chromatography on nitrilotriacetic acid (NTA)-agarose, followed by passage through SP Sepharose in intermediate salt buffer and binding to heparin-Sepharose (see Materials and Methods). The nearly homogeneous Gp44 preparation (Fig. 1C, lane 2) stimulated excision *in vitro* (Fig. 1D and E), demonstrating that Gp44 activates Int to catalyze *attL* × *attR* recombination. For some experiments, Gp44 was purified from an intein fusion, leaving the native polypeptide following thiol-induced cleavage. Where tested, intein-purified native Gp44 gave activities indistinguishable from those of the His-tagged version (Fig. 1E; also data not shown). As shown in Fig. 1F, Gp44 is active after heat denaturation at 80°C for 10 min.

Physical properties of Gp44. We confirmed the full-length nature of purified native Gp44 using matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS), which gave MH^+ values close to that expected for the 71-residue polypeptide following removal of the N-terminal methionine in *E. coli* (Fig. 2A). Gp44 is at least partially structured in solution based on its circular dichroism spectrum, which shows dips near 205 and 225 nm that are indicative of α-helical content (Fig. 2B). We were particularly interested in the oligomeric state of Gp44, since the integrase protein binds to *att* sites as dimers and forms tetrameric complexes during recombination.

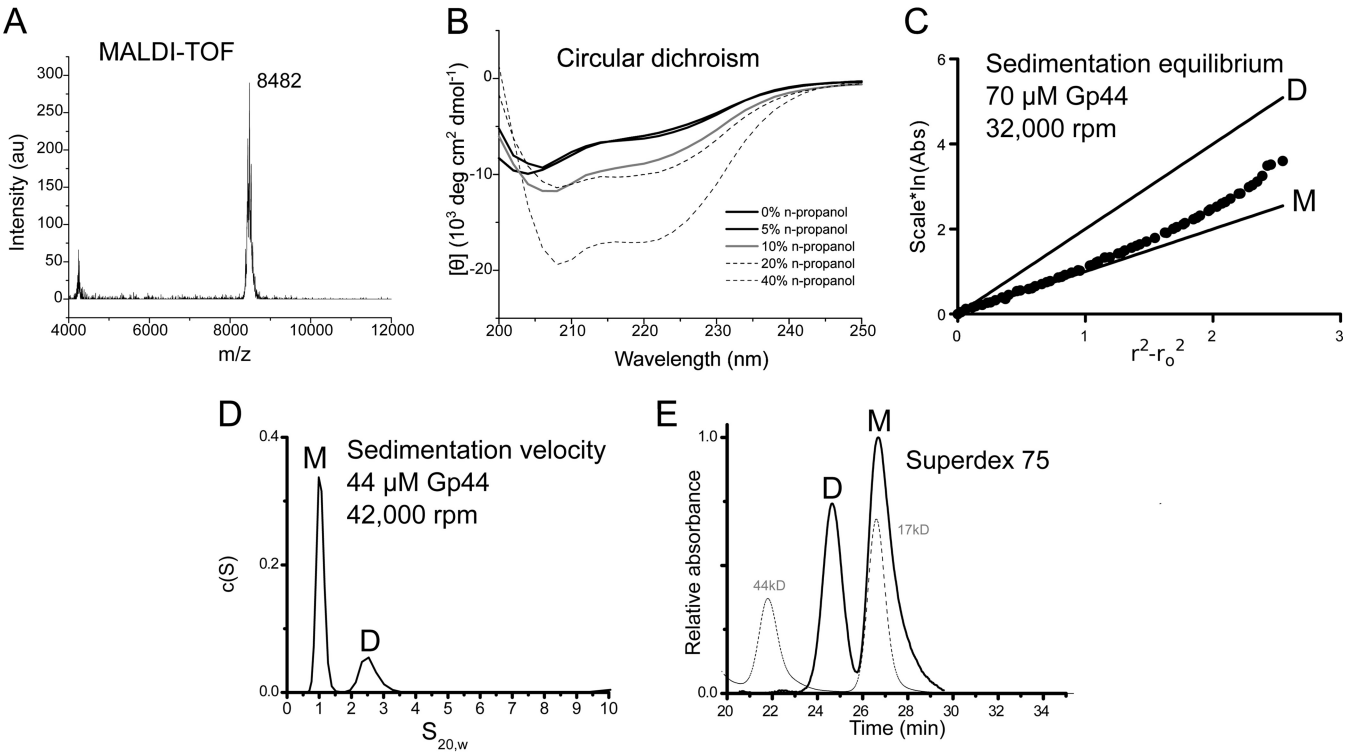

FIG 2 Properties of Gp44. (A) MALDI-TOF MS analysis of purified native Gp44. The calculated mass of the 71-residue polypeptide lacking an N-terminal methionine is 8,442 Da. (B) Circular dichroism of Gp44. Data were measured at 25°C in 50 mM PO_4 (pH 7.0) at a protein concentration of 22 μM and increasing concentrations of *n*-propanol to promote secondary structure. (C) Sedimentation equilibrium (SE) analysis was performed at three speeds and two concentrations at 4°C. Linearized data for a single speed and concentration are shown here, where scale = $2RT/[M(1 - \bar{\nu}\rho)\omega_2]$, where R is 1.98 kcal/mol, T is temperature, and M is molar mass. Other symbols are defined in Materials and Methods. M and D indicate the calculated data lines expected for single species monomer and dimer, respectively. The upwards curvature in the experimental data is indicative of an associating system. Global fit of the radial distributions gives an estimate of 13 μM for the dimer K_d (see Materials and Methods). (D) Sedimentation velocity (SV) analysis of Gp44. A 44 μM monomer was sedimented at 42,000 rpm at 4°C. Shown is the distribution of species derived from the fitting of the data with the Lamm equation. (E) Size exclusion chromatography on a Superdex 75 10/300 column at 25°C, where Gp44 was injected at 246 μM. Dashed gray lines represent standards. The analyses presented in panels C to E were performed in 20 mM Tris (pH 7.5), 300 mM NaCl.

Sedimentation equilibrium analysis at 70 to 120 μM indicates that Gp44 exists as monomers and dimers at this concentration, with an estimated dimer dissociation constant (K_d) of 13 μM (Fig. 2C; see also Materials and Methods). Sedimentation velocity (Fig. 2D) and size exclusion chromatography (SEC) (Fig. 2E) are consistent with this finding, with each method showing two species present. The apparent molecular weights of the monomer and dimer peaks from the sizing column relative to calibration standards are substantially higher than expected, indicating that Gp44 has a nonglobular structure in solution that increases its Stokes radius. Thus, Gp44 adopts an extended conformation and has the ability to self-interact, much like the coiled-coil domains of the integrase protein.

Gp44-activated excision reaction. Figure 3A and C show *attL* × *attR* (LR) excision reactions employing increasing amounts of Gp44 with a constant amount (50 nM) of Int. Yields of excision products increase sharply between 100 and 200 nM Gp44. LR recombination rates with saturating amounts of Gp44 are rapid and slightly higher than *attB* × *attP* (BP) recombination rates in the absence of Gp44 (Fig. 3D), but in both cases, yields of recombinant products plateau at about 75% of the substrate DNA. LR reactions with saturating amounts of Int, but subsaturating amounts of Gp44, generate lower rates and lower yields of excision products (Fig. 3D), implying there is little turnover of Gp44 to stimulate new reactions. An alternative explanation is that the plateau behavior reflects an equilibrium state generated by reactions back to the *attL* and *attR* parental forms under limiting Gp44 concentrations. Stoichiometric requirements of the RDF for the excision reaction by Bxb1 and ϕC31 integrases have also been reported (7, 8, 37).

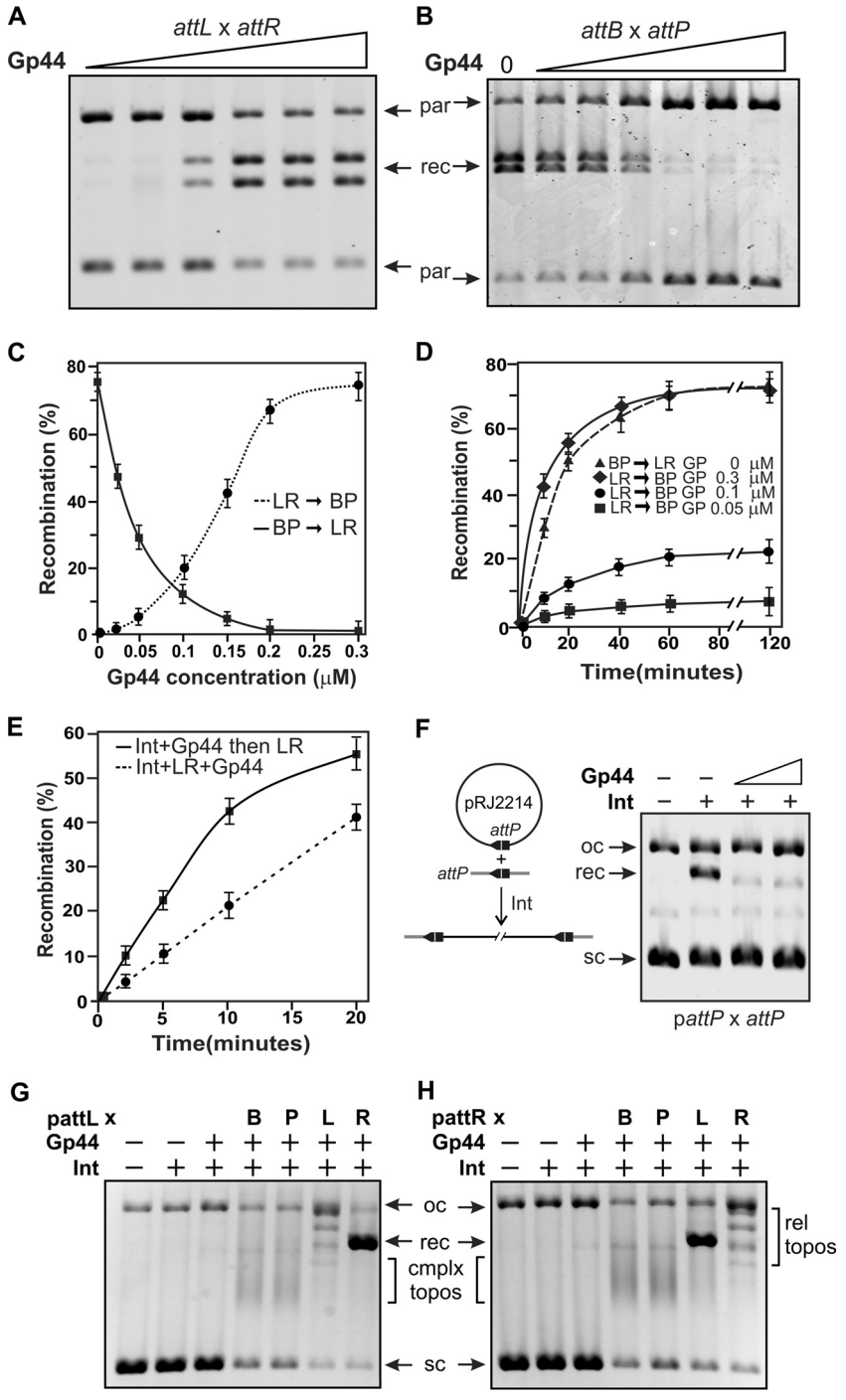

FIG 3 Properties and specificities of reactions with Gp44 *in vitro*. (A) LR recombination reactions on pRJ2986 with 50 nM Int and 25, 50, 100, 150, 200, and 300 nM Gp44. (B) BP recombination reactions on pRJ3244. The BP substrate (pRJ3244) is essentially identical to pRJ2986 (Fig. 1D), except the *attB* and *attP* sites are present in direct repeat configuration. (C) Plot of product yields after 60-min LR (dashed lines) and BP (solid lines) reactions performed with increasing amounts of Gp44. Error bars here and elsewhere indicate standard deviations. (D) Rates of LR and BP recombination reactions with different amounts of Gp44 as indicated. (E) Preincubation of Gp44 with Int accelerates the LR recombination reaction. Gp44 was incubated with Int for 10 min at 23°C and then added to the reaction buffer at 30°C containing pRJ3244 (solid lines), or Int and Gp44 were added sequentially to the reaction buffer at 30°C containing pRJ3244 (dashed lines). (F) Inhibition of *attP* × *attP* reactions by Gp44. On the left is a schematic of the intermolecular reaction in which pRJ2214 containing a single *attP* plus a 100-bp fragment containing *attP* was incubated with Int in the presence or absence of Gp44. The recombinant product is a linear DNA 100 bp longer than the plasmid. In the gel on the right, Int and Gp44 (100 and 200 nM) were added to the reaction as indicated. (G) pRJ2191 containing *attL* was incubated for 1 h with a 3-fold molar excess of

(Continued on next page)

There is a demonstrable effect by the order of addition of components to LR reactions. As shown in Fig. 3E, the reactions are most efficient when Int plus Gp44 are preincubated together in solution prior to the addition of substrate DNA. This feature has been observed for Gp3-activated LR reactions by the ϕC31 integrase (8) and is consistent with the RDF remodeling its cognate integrase into a more active conformation. When added, Gp44 was preincubated with Int for most of the experiments described in this paper, including the DNA binding reactions described below, which are also enhanced by preincubation.

Inhibition of BP and PP recombination by Gp44. Figure 3B and C illustrate the effects of increasing amounts of Gp44 on intramolecular BP integration reactions. Gp44 potently inhibits BP recombination, with significant reduction in the yield of products occurring with smaller amounts of Gp44 than that required for activation of LR excision. At a Gp44 concentration where LR recombinants are barely detectable (50 nM), BP recombination is reduced by >50%. At amounts of Gp44 that give maximal activation of LR excision, BP recombination is reduced to barely detectable levels. Gp44 also inhibits the weak *attP* × *attP* (PP) reaction catalyzed by the A118 Int (Fig. 3F) (23, 24).

Specificity of Gp44-regulated reactions. Previous studies showed that Int alone promoted only BP and weak PP recombination reactions (23, 24). We evaluated the effect of Gp44 on noncognate partner recombination using the more stringent intermolecular reactions between supercoiled plasmids containing *attL* or *attR* and linear DNA fragments containing each of the four *att* sites. The LR reactions with Gp44 generate linear recombinant products, as expected (Fig. 3G and H, lanes 7 and 6, respectively). Gp44-activated *attL* × *attL* and *attR* × *attR* reactions generate equivalent amounts of relaxed circular products (Fig. 3G and H, lanes 6 and 7, respectively) that migrate as supercoiled molecules in gels containing ethidium bromide (see Fig. S1 in the supplemental material), confirming they are covalently closed. Ethidium bromide gels also reveal that a small amount of covalently closed relaxed circular products are generated in LR reactions.

Formation of the covalently closed relaxed products can be explained by the formation of a synaptic complex whereby the *att* sites are aligned in the reverse (antiparallel) orientation. Cleavage and exchange of DNA strands occur, but ligation is prevented in the recombinant configuration due to the mismatch of core nucleotides (see references 34 and 38). Multiple subunit rotations that regenerate the parental *att* site sequences are required for ligation, which results in a loss of DNA supercoils. Relaxation of supercoiled plasmid substrates undergoing LL and RR reactions have also been reported for RDF-activated Bxb1, and recombinant products reflecting antiparallel synapses from LL and RR reactions with symmetric core sequences have been observed in RDF-activated Bxb1 and hyperactive ϕC31 integrase reactions (34, 38). Fully ligated relaxed circular products from processive DNA exchanges are also efficiently generated during A118 Int-catalyzed BP reactions on substrates in which the central 2 bp at the crossover sites are not equivalent (not shown). In addition, ethidium bromide gels reveal relaxed ligated circular products in wild-type BP reactions at early reaction times (not shown) that are consistent with the products from antiparallel synapses observed previously (23).

Gp44-activated integrase reactions between plasmids containing *attL* or *attR* and linear fragments containing *attB* or *attP* also generated a mixture of complex plasmid topoisomers migrating between the supercoiled and relaxed species (Fig. 3G and H). The use of radiolabeled linear substrates confirmed that these are not recombinant

FIG 3 Legend (Continued)
100-bp fragments containing *attB*, *attP*, *attL*, and *attR* plus Int and Gp44 as indicated. (H) Same as for panel G except pRJ2193 containing *attR* was the plasmid substrate. The samples in panels G and H were incubated with SDS and proteinase K before electrophoresis. The linear recombinant (rec) products, regions exhibiting relaxed plasmid topoisomers (rel topos), complex (cmplx) topologically linked plasmid products, and open circular (oc) and supercoiled (sc) plasmid forms are denoted. See Fig. S1 for reactions similar to those in panels G and H, except that electrophoresis was performed in the presence of ethidium bromide.

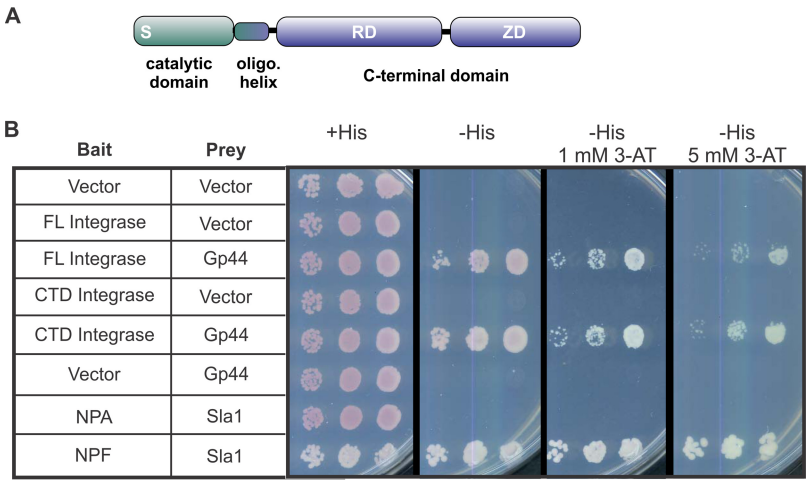

FIG 4 Yeast two-hybrid analyses showing that Gp44 interacts with the Int CTD. (A) Domain structure of the A118 integrase showing the catalytic domain with the active site serine (S), oligomerization (oligo.) helix E, and C-terminal domain (CTD) that binds *att* DNA. (B) Yeast two-hybrid assays for interactions between Gp44 and full-length (FL) or CTD Int. Serial 10-fold dilutions of yeast strains containing fusions to the GAL4 DNA binding domain (bait) and GAL4 activation domain (prey) plasmids were spotted onto synthetic medium with or without histidine and 3-amino-1,2,4-triazole (3-AT), as indicated. Yeast proteins Sla1 × NPF and Sla1 × NPA are included as positive and negative controls, respectively (49).

products (not shown), even though their formation depends on both substrates in the reaction. Their relative migration is unaltered upon enzymatic nicking (not shown), and they run in a similar relative position in gels with ethidium bromide (Fig. S1), consistent with most being complex topologically interlinked (probably open circular knotted) forms. We conclude that Gp44 enables Int to engage *attL* or *attR* in a reactive synaptic complex with *attB* or *attP* that does not yield recombinant product, presumably because the *att* sites are synapsed in an antiparallel orientation.

Gp44 and integrase form complexes *in vivo*. Yeast two-hybrid assays (39) were initially employed to test whether there was a specific interaction between Gp44 and Int. In the experiment shown in Fig. 4B, integrase was fused to the GAL4 DNA binding domain, and Gp44 was fused to the GAL4 activation domain. The *Saccharomyces cerevisiae* reporter strain harboring these two plasmids gave strong growth in the absence of histidine and in the presence of the HIS3 enzyme competitive inhibitor 3-aminotriazole, providing evidence that Gp44 and Int form a complex. The monomeric C-terminal domain (CTD; residues 148 to 452) (Fig. 4A) (23, 24) was substituted for the full-length integrase and was found to also be sufficient for interaction with Gp44 (Fig. 4B). These results provide evidence that Gp44 binds to the CTD of the A118 integrase.

Complexes formed between Gp44 and integrase on DNA *in vitro*. We have been unable to detect binding of Gp44 to DNA fragments containing the different *att* sites in electrophoretic mobility shift assay (EMSA) experiments lacking integrase (Fig. 5A; also data not shown). However, the presence of Gp44 alters the electrophoretic migration of all four of the Int-*att* complexes in an *att*-specific manner (Fig. 5B to E). Gp44 plus Int form strongly supershifted complexes on *attL* and *attR*, but the complexes appear unstable, as evidenced by the smear down to the Int-only band even at high Gp44 concentrations. Gp44-bound Int-*attB* complexes migrate faster than those containing only Int. On gels with lower bisacrylamide concentration, a slightly slower migrating intermediate complex is evident at low Gp44 concentration (Fig. 5D, bottom); we interpret this profile to imply that Gp44 bound to one Int subunit slightly decreases electrophoretic mobility, whereas Gp44 bound to both Int subunits increases the mobility of the *attB*-Int dimer complex. Gp44 bound Int-*attP* complexes are slightly supershifted relative to Int-*attP* complexes. Slow electrophoresis on low-bisacrylamide gels resolves an intermediate complex, suggesting two units of Gp44 are sequentially bound to the *attP*-Int dimer complex (Fig. 5E, bottom). The differences in the electro-

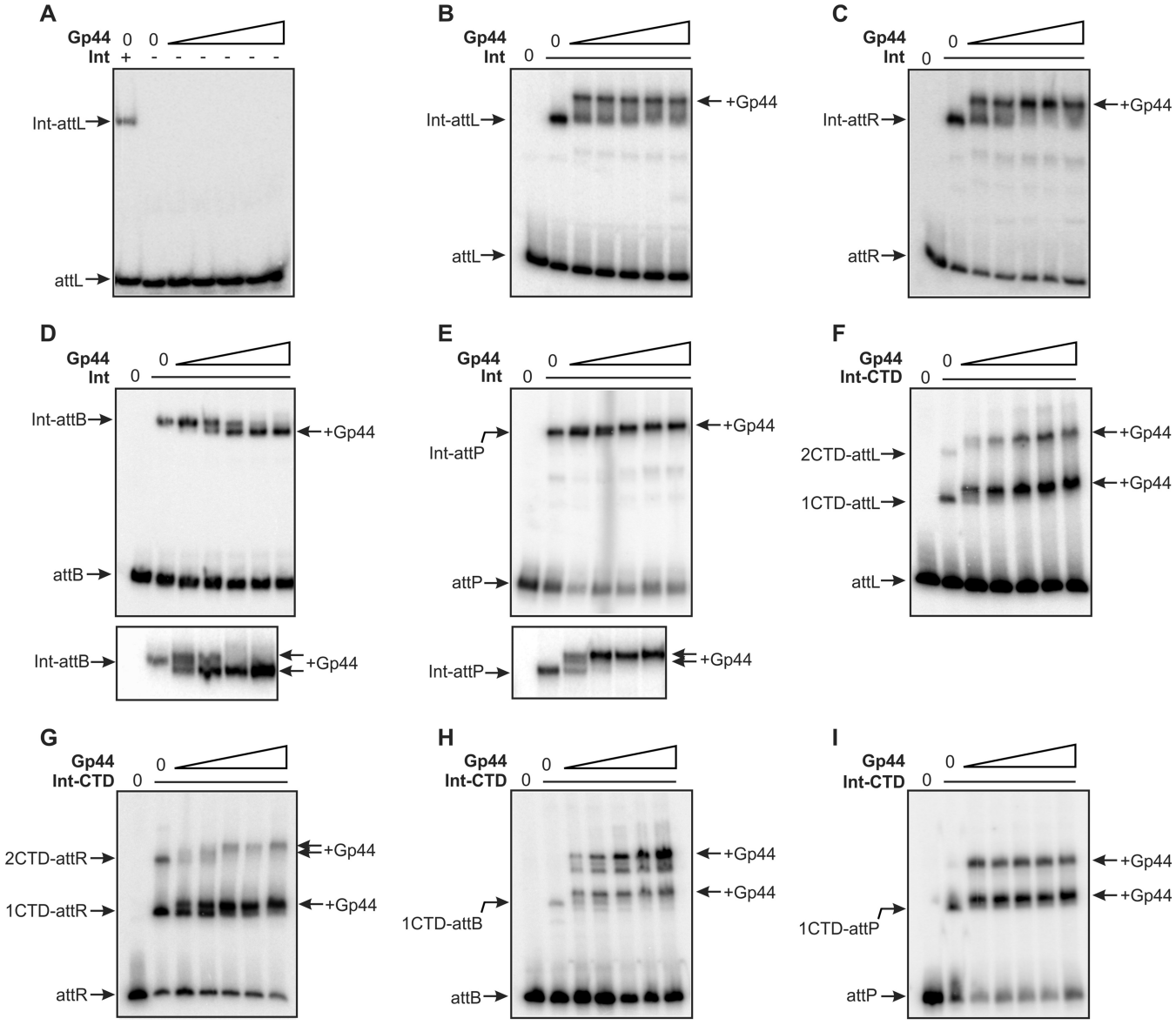

FIG 5 Gp44-Int-*att* complexes *in vitro*. (A) EMSA of Gp44 binding to *attL*. Gp44 was added to DNA in 2-fold increasing concentrations from 6.25 to 100 nM. Int was only added to the reaction in the left lane. No direct binding to DNA containing *attL* (or *attR*, *attB*, or *attP* [data not shown]) is detectable. (B to E) Gp44 was preincubated with Int prior to incubation with *attL*, *attR*, *attB*, and *attP* probes, respectively. Final Gp44 concentrations varied by 2-fold increasing amounts from 6.25 to 100 nM with 5 nM Int. The locations of complexes containing Int or Gp44 plus Int bound to *att* DNA are marked. Bottom images in panels D and E show complexes after slow electrophoresis through 59:1 acrylamide:bisacrylamide gels. (F to I) Gp44 complexes formed with Int-CTD on *attL*, *attR*, *attB*, and *attP* DNA, respectively. Final protein concentrations were 20 to 320 nM Gp44 in 2-fold increasing amounts with 50 nM Int-CTD on *attR* and *attP* and 80 to 1,200 nM Gp44 in 2-fold increasing amounts with 200 nM Int-CTD on *attL* and *attB*. The locations of complexes bound by I and 2 Int-CTDs alone and with Gp44 are noted.

phoretic migrations of the Gp44-Int complexes on the various *att* sites are consistent with Gp44 altering the trajectory of the CC motifs to form structural architectures that are *att* site specific (see Discussion).

We showed previously that the monomeric Int-CTD binds with very different affinities to the four individual *att* half sites and noncooperatively to *att* full sites, unlike full-length Int (23). We asked how Gp44 influences Int-CTD binding to the different *att* full sites. The presence of Gp44 results in slower migrating complexes containing one and two Int-CTD bound to *attL* and *attR* (Fig. 5F and G). On *attR*, two discrete Gp44-dependent species can be resolved on the two Int-CTD complex (see also Fig. S2), implying one Gp44 unit bound per Int-CTD. The presence of Gp44 enhances binding of the two Int-CTD complexes on *attB* and *attP*, which in the absence of Gp44 are hardly

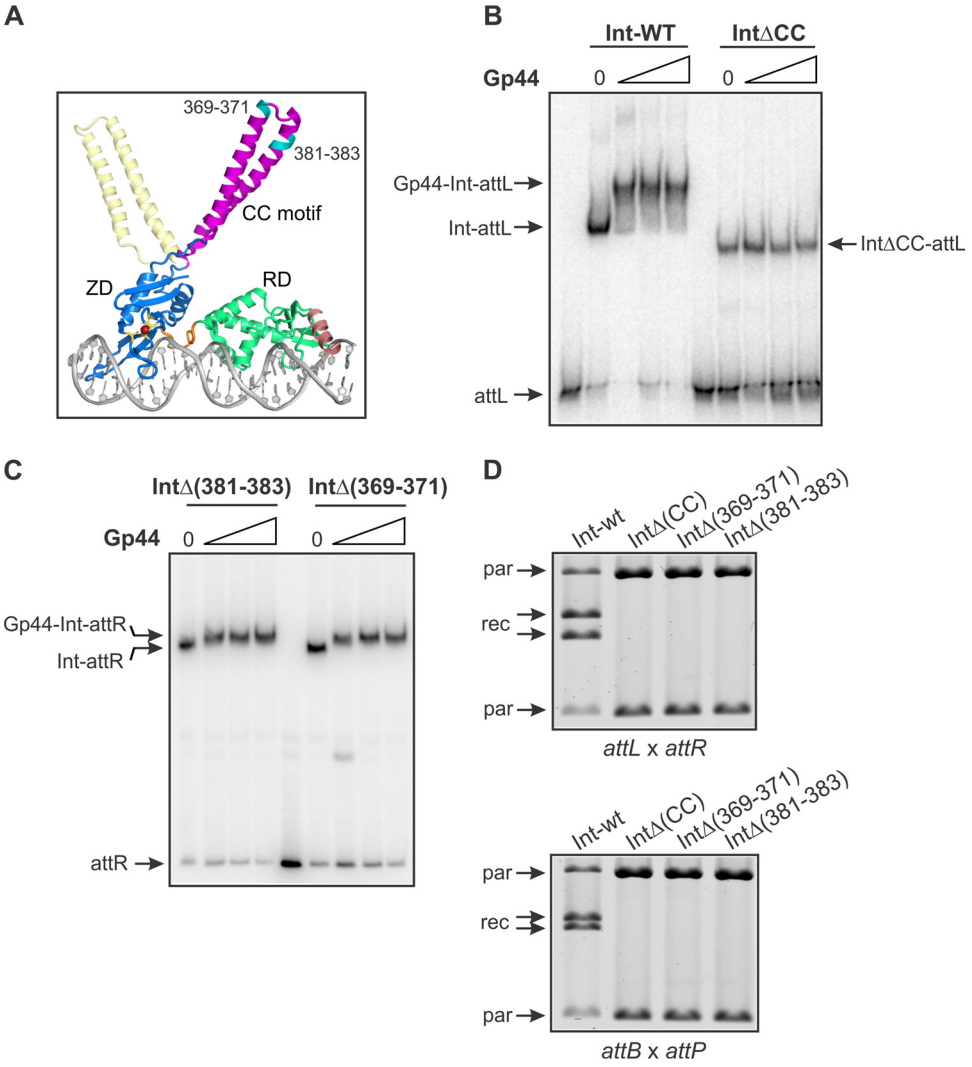

FIG 6 Gp44 binding to Int mutants. (A) Structure of the L1 Int-CTD bound to the *attP* half site (Protein Data Bank [PDB] accession no. 4KIS) highlighting the mobile CC motifs (magenta). The CTD including the CC motif is from molecule C with the tip from molecule A appended. The CC motif in transparent yellow is aligned from molecule A to illustrate another trajectory of the CC motif emanating from the zinc domain (ZD; blue). The linker peptide between the recombinase domain (RD; green) and ZD is colored orange, and the helix E region that connects the CTD to the N-terminal catalytic domain is in salmon. The locations of two deletions near the tip of the CC motif analyzed in this work are in cyan. The ZD recognition sequence is proposed to be 5 bp closer to the RD in the *attB* half site, which will position the ZD with its protruding CC motif on the opposite side of the DNA helix. The amino acid sequences of Int proteins from related phages A118 and L1 are 98% identical, and A118 Gp44 binds to and activates excision by L1 Int. (B) Gp44 does not form a stable complex with Int missing the CC motif. EMSAs showing binding reactions on *attL* DNA with 2-fold increasing concentrations of Gp44 beginning with 25 nM and 5 nM Int-wt or IntΔCC. (C) Gp44 forms stable complexes with Int mutants containing deletions near the tip of the CC motif. Gp44 and Int concentrations were the same as for panel B, and the DNA probe was *attR*. (D) Int deletion mutants are defective for Gp44-activated LR (top, 200 nM Gp44) and BP (bottom, no Gp44) recombination. Int proteins were present at 50 nM, and reactions were for 60 min.

evident at this Int-CTD concentration (Fig. 5H and I; see also Fig. S3, left panel). Interestingly, the migration of the two Int-CTD complex on *attB* is retarded when bound by Gp44 (Fig. 5H), whereas the full-length Int dimer complex on *attB* exhibits a slightly faster migration when bound by Gp44 (Fig. 5D).

Gp44 binding requires the integrase CC motif. Rutherford et al. proposed that binding of the RDF to serine integrases may reorient the trajectories of the mobile CC motifs on the Zn-ribbon subdomain of the CTD (Fig. 6A) (24, 28). This model suggests that Gp44 may interact with the CC motifs. To explore this possibility, Gp44 binding to an Int mutant with the entire CC motif deleted was evaluated. Gp44 was unable to

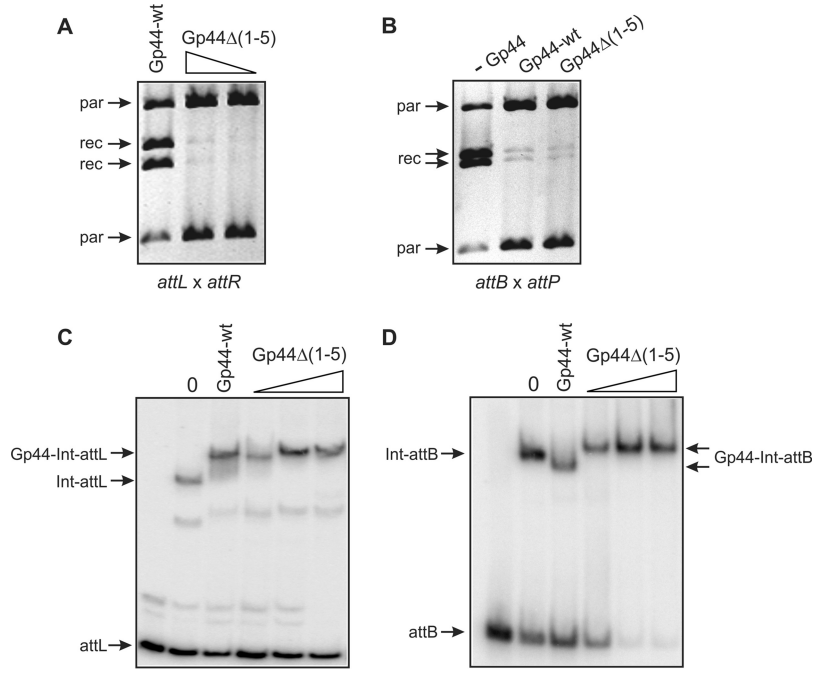

FIG 7 Gp44 separation-of-function mutant. (A) Gp44Δ(1–5) is defective for promoting LR recombination. LR recombination was performed with 50 nM Int and either 200 nM Gp44-wt or 200 and 400 nM Gp44Δ(1–5). (B) Gp44Δ(1–5) is active for inhibiting BP recombination. BP recombination was performed without Gp44 or in the presence of 200 nM Gp44-wt or Gp44Δ(1–5). (C and D) EMSA complexes formed by Gp44Δ(1–5) and Int on *attL* and *attB* DNA, respectively. Gp44-wt (100 nM) or Gp44Δ(1–5) (2-fold increasing concentrations beginning with 25 nM) were incubated with 5 nM Int and *att* DNA. Lane 0 is Int without Gp44.

detectably form a complex with *attL* (Fig. 6B) or with the three other *att* complexes bound by Int with a deleted CC motif (IntΔCC) (data not shown). On the other hand, Gp44 bound normally to two Int mutants containing three-residue deletions near the tip of the CC motif (Fig. 6A and C). Each of these Int mutants is defective in promoting BP or Gp44-activated LR recombination (Fig. 6D). We conclude that at least part of the Gp44 binding site on Int involves the CC motif but not residues at its tip that are important for recombination.

Gp44 separation-of-function mutant. During the course of this work, we found that a Gp44 mutant containing a short truncation at its N terminus abolished activation of LR recombination (Fig. 7A) but retained high-affinity binding to Int-*att* complexes (Fig. 7C and D). Remarkably, Gp44 with a deletion at residues 1 to 5 [Gp44Δ(1–5)] binding to Int-*attB* complexes generates slower mobility complexes than wild-type Gp44, which forms faster migrating complexes (Fig. 7D). In addition, Gp44Δ(1–5) binding to Int-*attL* complexes appear more stable than wild-type Gp44 (Gp44-wt) (Fig. 7C). Moreover, Gp44Δ(1–5) supershifts one and two Int-CTD complexes on *attP*, but the 2-Int-CTD complex exhibits a markedly greater supershift when bound by Gp44Δ(1–5) than by Gp44-wt (Fig. S3). These data are consistent with the N-terminal region of Gp44 modulating the trajectory of the Int CC motif, which affects the electrophoretic migration of the complex. The properties of Gp44Δ(1–5) demonstrate that binding to Int alone is insufficient for activation of LR recombination and provide evidence that two separate regions on Gp44, and by inference Int, may be required. On the other hand, Gp44Δ(1–5) remains functional for inhibiting BP recombination (Fig. 7B). Thus, the requirements for activation of LR and inhibition of BP recombination by Gp44 are distinct.

DISCUSSION

We show here that the *L. monocytogenes* phage A118-encoded protein Gp44 functions as a recombination directionality factor (RDF) for the phage integrase.

Whereas no detectable *attL* × *attR* excision reaction is catalyzed by Int alone, the presence of Gp44 promotes excision with rates slightly higher than that of the *attB* × *attP* integration reaction in its absence. The presence of Gp44 also strongly inhibits the BP integration reaction. Thus, Gp44 expression levels must be extremely low upon infection in order for the phage to generate a stable prophage within the *comX* gene in the *Listeria* chromosome but must be expressed when the lytic cycle is triggered. Since migration of *Listeria* from the mammalian cell phagosome into the cytoplasm requires precise excision of the phage to restore the *comX* coding region (27), Gp44 expression must also be induced during host infection. *gp44* is located within the cluster of phage A118 genes involved in the control of lysogeny but is separated from the Int gene by 16 open reading frames (ORFs) (Fig. 1B). It is likely that Gp44 has additional roles in A118 biology, as has been described for other RDFs (9, 40–42).

Currently identified RDFs for serine integrases fall into two size classes (1). The three small RDFs (A118 Gp44, ϕRv1 Rv1584c, and TP901-1 orf7) range from 64 to 73 amino acid residues and are relatively basic (pI 9.3 to 10.7). However, there is little resemblance between these proteins at the primary sequence level. The large RDFs, exemplified by Gp47 from phage Bxb1 and the related Gp3 proteins from *Streptomyces* phages ϕC31 and ϕBT1, are 244 to 255 amino acid residues in length, but these and the small RDFs are unrelated in sequence. The small RDFs, as shown in this work with Gp44, exhibit functional properties similar to those of the large RDFs, and members of both groups appear to function exclusively through binding to their respective integrases. The mechanism of action for serine recombinase RDFs is unrelated to the RDFs (or Xis proteins) that control tyrosine recombinases and function primarily through binding and changing DNA architecture.

Surprisingly, Gp44 enables A118 Int to assemble a diversity of chemically active synaptic complexes, but only LxR reactions generate recombinant products. Gp44-stimulated synaptic complexes that do not yield recombinant products include LxL and RxR, both robustly formed but in the antiparallel orientation; and LxB, LxP, RxB, RxP, and LxR, which are formed less efficiently but also presumably in the antiparallel orientation. These are all chemically active, as evidenced by resulting topological changes in the substrate DNA (Fig. 3G and H; see Fig. S1 in the supplemental material), but they do not generate recombinant products presumably because the core (crossover) nucleotides are not complementary after odd numbers of subunit rotations, which prevents ligation. The ability of the RDF to enhance a more promiscuous reaction seems to be a general property of serine integrases, as it has been also documented in the Bxb1, ϕRv1, ϕC31, and ϕBt1 systems (8, 9, 38). As elaborated below, the RDFs may remodel Int to both disrupt an inhibitory conformation and to directly promote specific synaptic complex architectures by orienting the trajectory of the CC motif. After formation of these synaptic complexes, the check with respect to whether they can generate recombinant product is solely at the ligation step after DNA exchange by subunit rotation (1, 38).

Although we see evidence for dimerization at high concentrations, Gp44 is likely to be primarily monomeric at solution concentrations active for Int-DNA binding (~10 nM) (Fig. 5) or excision (half maximal stimulation occurs at ~150 nM under standard conditions with 50 nM Int monomer) (Fig. 3C). Our estimated dimerization K_d (13 μM) and the large monomeric fraction observed by gel filtration (Fig. 2E) are consistent with this assumption. The weak self-association of Gp44 could still play a role in the context of Int-Gp44-*attL* or Int-Gp44-*attR* complexes, since multiple copies of the RDF would be juxtaposed within an LR synaptic complex. Gp44 is fully active after heat denaturation (Fig. 1F), and CD analysis indicates that a native solution conformation with partial helical structure can adopt a highly helical structure (Fig. 2B). Given its nonglobular shape, it seems likely that Gp44 is partially unstructured in solution but may adopt a specific conformation upon binding to Int. A Gp44 dimeric structure could also be relevant in additional functions it may perform during lytic growth.

Gp44 functions stoichiometrically with Int to activate LR recombination, as evidenced by the lack of apparent turnover to activate additional Int complexes when

Gp44 levels are limiting (Fig. 3D). Int binding experiments imply that all four Int subunits are bound by Gp44 under conditions favoring excision (Fig. 5). Moreover, two distinct species whose electrophoretic mobility is shifted by increasing Gp44 have been resolved on Int dimer complexes formed on *attB* (Fig. 5D) and *attP* (Fig. 5E). On *attR*, the DNA-bound complex containing two Int-CTDs forms two distinct Gp44-shifted species (Fig. 5G and S2). Interestingly, BP recombination is efficiently inhibited at considerably lower concentrations of Gp44 than are required for stimulation of the LR reaction (Fig. 3C). One explanation is that fewer Gp44 molecules than Int molecules are required for BP inhibition, suggesting that Gp44 can function substoichiometrically in this case. Alternatively, higher levels of Gp44 may be required to unlock competitive intramolecular coiled-coil interactions in the LR reaction that are not present in the Int-bound *attP* and *attB* sites. Where evaluated, the large RDFs for serine integrases appear to behave similarly with respect to relative stoichiometries for LR excision versus BP inhibition (7–9). However, the various RDFs behave differently with respect to their apparent affinities for and resulting electrophoretic migration changes to the four *att*-Int complexes. Indeed, the increased mobility induced by Gp44 binding to *attB-Int* complexes we observed (Fig. 5D) has not previously been reported for the large RDFs.

The Int dimer can probably adopt multiple conformations in solution, as the catalytic and two DNA binding domains are separated by polypeptide linkers that may be flexible when Int is not bound to *att* sites (2). Preincubation of Gp44 with Int prior to the addition of DNA enhances LR recombination rates (Fig. 3E), indicating that Gp44 interacts with Int in solution, probably to both promote conformations that are more competent for binding DNA as well as for LR synapsis. Specifically, binding of Int by Gp44 may discourage interactions between CC motifs on each subunit of the Int dimer, which otherwise may conform the Int dimer into a structure that binds DNA poorly (24).

Gp44 association with Int causes an apparent change to the structure of the DNA complex in an *att*-specific manner that is revealed by native gel electrophoresis. The migration changes are consistent with reorientation of the CC motifs upon Gp44 binding, as proposed by Rutherford et al. (24, 28). Int complexes with *attP*, *attB*, and *attL* or *attR* are modeled to have different structures due to the unique positions of the ZD relative to the catalytic and RD domains when bound to the different *att* sites. Changing the trajectories of the CC motifs by Gp44 would be expected to have different relative effects on the migration of the DNA complex through polyacrylamide, reminiscent of the effects of coiled-coil regions on basic leucine zipper domain (bZIP) protein-DNA complexes (43). In the case of *attB*, the intermediate complex formed with smaller amounts of Gp44 is slightly retarded in its migration, but the final complex formed with greater amounts of Gp44 migrates faster than the Int-only *attB* complex. These Gp44-induced changes in CC trajectories could be the result of Gp44 stabilizing a single CC conformation versus the ensemble of CC orientations that are likely to be present in the absence of Gp44.

An alternative explanation for the changes in electrophoretic migrations is that Gp44 binding has an effect on DNA curvature of the Int-bound complex. We consider this unlikely because no evidence of DNA bending has been observed in circular permutation gel electrophoresis assays (44) of *attP*-containing fragments bound by Int or Int plus Gp44 (S. Mandali and R. C. Johnson, unpublished data). Moreover, the DNA within the X-ray structure of L1 Int-CTD bound to a half site exhibits little overall curvature (24).

Int mutations removing the entire CC motif abolish Gp44 binding, but mutations near the distal tip that also eliminate *in vitro* recombination do not affect Gp44 binding, indicating that a segment closer to the base of the CC motif (Fig. 6A) is required. An additional region on Int may also be involved because Gp44 mutants that are missing N-terminal residues are still able to efficiently bind Int but cannot activate LR recombination. Significantly, Int dimers containing Gp44Δ(1–5) complexes on some *att* sites migrate differently than those with Gp44-wt (Fig. 7D; see also Fig. S3). We propose that Gp44 associates primarily with the base-stem region of the CC motif and also interacts with an additional region on Int. The additional region could be on the body of the ZD,

in which case the CC trajectories relative to the ZD would be expected to be similar for P, P′, B, and B′ half sites. Alternatively, if the additional region involves the RD, the trajectories relative to the ZD would likely be quite different for P and P′ versus B and B′, since the recombinase and zinc ribbon domains are on the same face of the DNA for P and P′ but are on opposite faces of the DNA for B and B′. It is also possible that the N terminus of Gp44 plays a different role, such as mediating RDF-RDF interactions.

Whatever the specific contacts made by the Gp44 N terminus, a two-contact model for Gp44 would provide a mechanism whereby it could control the trajectories of the CC motifs to favor LR synapsis and recombination. The ability of Gp44Δ(1–5) to efficiently inhibit BP recombination implies that an interaction with only a single region on Int, presumably on the CC motif, is sufficient to discourage BP synapsis and recombination, whereas LR excision requires additional RDF-mediated contacts.

MATERIALS AND METHODS

Strains, phages, and recombinant plasmids. A list of strains and plasmids used in this work is given in Table 1. The A118 *attL* and *attR* sites were PCR amplified from DP-L3689 (*L. monocytogenes* strain 10403 lysogenized with A118, provided by R. Calendar, University of California-Berkeley) and initially cloned between the EcoRI and SalI sites of pBR322 to give pRJ2191 and pRJ2193, respectively. Integrase was expressed from the pET15b derivative pRJ2184 (23). Int deletions (Table 1) were generated in pRJ2184 by QuikChange methods. pRJ2986, the plasmid substrate for the LR deletion reaction *in vitro*, was derived from pBR322 and contains *attL* between the EcoRI and HindIII sites, *attR* in directly repeated orientation with respect to *attL* between the BamHI and SalI sites, and *lacPOZ* inserted into the EcoRV site. pRJ3244 is identical to pRJ2986 except that it contains *attB* and *attP* instead of *attL* and *attR*, respectively, in direct repeat orientation.

The reporter for *in vivo* recombination, pRJ3225, was constructed by amplifying the *attL-lacZ-attR* region from pRJ2986 and inserting the fragment into the SmaI site of pMS421 (pSC101^{ori}; M. Susskind, University of Southern California). Recombination by the A118 integrase will delete the *lacZ* gene to generate white colonies on X-Gal medium. pRJ3245 that supplies Int is derived from pACYC184 and contains the *lacPuv5* promoter cloned between SalI and BamHI and *int* between BamHI and XbaI. For identification of the RDF, segments of the coding sequences from the regulatory genes *gp32* to *gp45* were PCR amplified from phage A118 DNA (gift from R. Calendar) and cloned between the PstI and SalI sites of pUC18 (Table 1). The coding region for Gp44 with a 6His tag at its N terminus was PCR amplified and inserted into pET11a between NdeI and BamHI to give pRJ2990.

Protein purification. Full-length and Int mutants were purified as described previously (23). ^{His}Gp44 was expressed in RJ3387 containing pRJ2990 (Table 1). One-liter LB cultures were grown at 37°C and induced for 1 h with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) when the optical density at 600 nm (OD_{600}) reached 0.6. Harvested cells were resuspended in 20 ml of lysis buffer (50 mM HEPES [pH 7.5], 200 mM NaCl, 20 mM imidazole, 10% glycerol, 1 mM β-mercaptoethanol, and 0.1 mM phenylmethylsulfonyl fluoride [PMSF]) and lysed by a French press. The lysate was centrifuged at 26,000 × *g* for 30 min, and the supernatant loaded onto a 1-ml Ni-NTA–agarose (MCLAB) column that was then washed with 20 mM HEPES (pH 7.5), 0.5 M NaCl, 40 mM imidazole, 10% glycerol, and bound protein eluted with the same buffer but with 300 mM imidazole and 0.2 M NaCl. Fractions containing ^{His}Gp44 were pooled, dialyzed into buffer A (20 mM HEPES [pH 7.5], 10% glycerol, 1 mM dithiothreitol [DTT], 0.1 mM EDTA) plus 0.3 M NaCl, and passed through SP Sepharose Fast Flow (GE Healthcare) in the same buffer. ^{His}Gp44 was then bound to heparin-Sepharose 6 Fast Flow (GE Healthcare) in the same buffer, and near-homogeneous fractions were eluted with buffer A containing 1 M NaCl. Gp44 preparations were typically stored in 20 mM HEPES (pH 7.5), 1 M NaCl, 1 mM DTT, 0.1 mM EDTA, and 40% glycerol at −20°C.

Native Gp44 was expressed from a pETDuet (Novagen)-derived vector in BL21(DE3) cells at 37°C. Gp44 was subcloned into the vector in-frame with a C-terminal Mxe intein (New England BioLabs) containing chitin binding domain and 6His affinity tags (pGV2710). Proteins were purified using nickel-nitrilotriacetic acid (Qiagen) and chitin (New England BioLabs) resins. Fusion protein was released by intein cleavage in 100 mM 2-mercaptoethanol overnight at 4°C. Preparations of Gp44 were further purified using SP Sepharose chromatography (GE Healthcare). Protein in 20 mM Tris (pH 7.5), 300 mM NaCl, and 1 mM DTT was concentrated at 4°C in a YM-3 Centricon (Millipore), and aliquots were flash-frozen in 10% glycerol with liquid nitrogen for storage at −80°C.

SE and SV analytical ultracentrifugation. Sedimentation equilibrium (SE) experiments were performed at 4°C with an XL-A analytical ultracentrifuge (Beckman Coulter) and a TiAn60 rotor with six-channel charcoal-filled Epon centerpieces and quartz windows. Radial absorption scan data at 280 nm for 70 and 118 μM Gp44 were measured at 18 and 20 h for each of three different rotor speeds (ω; 28,000, 30,000, and 32,000 rpm). Comparison of radial absorption scans verified that equilibrium had been reached. Data were analyzed using the programs SEDFIT (45) and SEDPHAT (46). Global fits to the SE radial distributions gave a dimer K_d value of 13 μM, with an estimated error of 0.01 μM based on a 1,000-iteration Monte Carlo simulation, as implemented in SEDPHAT. Sedimentation velocity (SV) experiments were performed at 4°C with two-channel charcoal-filled Epon centerpieces and quartz windows. Complete sedimentation velocity profiles were collected every 30 s at 42,000 rpm, followed by fitting to the Lamm equation using the program SEDFIT. For all centrifugation analyses, the partial specific volume

($\bar{v}$), solvent density (ρ), and viscosity (η) were derived from chemical composition by SEDNTERP (47). All centrifugation analyses were performed in 20 mM Tris (pH 7.5) and 300 mM NaCl.

SEC. Gel filtration experiments were performed at room temperature using a Superdex 75 10/300 column equilibrated in 10 mM Na·KPO_4 and 150 mM NaCl at 0.5 ml/min. A 20-μl sample of 246 μM Gp44 was injected and eluted isocratically and the eluent monitored using UV absorbance at 280 nm. The column was calibrated using the following proteins (Bio-Rad): thyroglobulin (670 kDa; Stokes radius [R_S] = 85 Å), γ-globulin (158 kDa; R_S = 52.2 Å), ovalbumin (44 kDa; R_S = 30.5 Å), myoglobin (17 kDa; R_S = 20.8 Å), and vitamin B_{12} (1,350 Da). Blue dextran (Sigma) was used to define the void volume of the column.

CD. Circular dichroism (CD) spectra were recorded using an Aviv CD spectropolarimeter with a 1-mm path length at 25°C. Data were recorded using a 1-nm bandwidth with 30-s averaging time and 2-nm steps. Spectra were normalized to the mean residue ellipticity using a protein concentration determined by absorbance at 280 nm. Gp44 was at 22 μM concentration in 50 mM Na·KPO_4 (pH 7.0), with 0 to 40% *n*-propanol in all measurements.

MALDI-TOF MS. MALDI-TOF mass spectrometry to determine the mass of the peptide was performed with a Bruker Microflex LHF spectrometer. Samples were desalted using a C_{18} ZipTip (Millipore), as per the manufacturer's instructions, and eluted in 100% acetonitrile (ACN). The matrix used was a solution of 1:1 2-(4-hydroxyphenylazo)benzoic acid (HABA) and α-cyano-4-hydroxy-cinnamic acid (CHCA) in 100% acetonitrile–0.1% trifluoroacetic acid (TFA). Data analyses were performed using the Bruker Daltonics flexAnalysis software version 3.3, and measurements were calibrated to a bovine RNase A standard (New England BioLabs) with a monoisotopic mass of 13,681.32 Da.

***In vitro* recombination reactions.** Standard recombination reactions were performed in 20 μl of buffer containing 20 mM HEPES (pH 7.5), 100 mM NaCl or KCl, 5 mM spermidine, 2.5 mM DTT, 5 mM $MgCl_2$, 30 μg/ml bovine serum albumin (BSA), and 5% glycerol. In most experiments, Gp44 and Int were preincubated together for 10 min at 23°C at 10 times the final concentration in reaction buffer without DNA. A 1/10 volume was then added to the reaction buffer containing substrate DNA at 30°C to initiate the reaction.

Typical reactions were performed with 1 pmol integrase and 4 pmol Gp44. Intramolecular deletion reactions were performed with 0.02 pmol plasmid DNA, and intermolecular integration reactions utilized 0.03 pmol supercoiled plasmid DNA and 0.09 pmol linear DNA (100-bp fragment derived from PCR). The reactions proceeded at 30°C and were terminated by inactivation at 65°C or addition of SDS. Intramolecular recombination reactions were digested with NdeI to generate linear deletion products prior to agarose gel electrophoresis.

***In vitro* DNA binding assays.** ^{32}P-labeled 100-bp DNA probes with centrally located *att* sites were generated by PCR after labeling one of the primers with γ-^{32}P-ATP using polynucleotide kinase and purifying by PAGE gels. In most experiments, various concentrations of Gp44 were incubated with a fixed concentration of integrase (50 nM) in binding buffer (20 mM HEPES [pH 7.5], 100 mM NaCl, 0.1 mg/ml BSA, 5 mM DTT, 5% glycerol, and 50 μg/ml sonicated salmon sperm DNA) for 10 min at 23°C. A 1/10 volume (2 μl) was then added to 18 μl of binding buffer containing approximately 0.1 pmol ^{32}P-labeled *att* probe. The reaction mixtures were incubated at 30°C for 30 min, and the protein-DNA complexes were separated on nondenaturing 6% polyacrylamide gels (37.5:1 acrylamide-bisacrylamide) containing 2% glycerol in 0.5× Tris-borate-EDTA (TBE) buffer at 4°C. Where noted, 59:1 acrylamide-bisacrylamide gels were used. Gels were dried and analyzed by phosphorimaging.

***In vivo* recombination assay.** RDF-activated *in vivo* excision was evaluated by transformation of pUC18 plasmids containing A118 genes into *E. coli* DH5α containing pRJ3225 (LR recombination reporter plasmid; pSC101^{ori} *str lacI*^{Q}) and pRJ3245 (Int^{+} p15A^{ori} *cat*) (Table 1). White colonies on LB plates containing ampicillin, streptomycin, chloramphenicol, 0.1 mM IPTG, and 20 μg/ml X-Gal indicated loss of *lacZ* by LR recombination. Restriction analysis of products from white colonies generated in the presence of *gp44* confirmed that the pRJ3225 plasmids had undergone the expected site-specific deletion.

Yeast two-hybrid analysis. *S. cerevisiae* strain PJ69-4A (48) was cotransformed with derivatives of yeast two-hybrid vectors pGAD GH (LEU2, GAL4 activation domain; Clontech) and pGBKT7 (TRP1, GAL4 DNA binding domain; Clontech) and examined for growth on plates lacking histidine. The entire coding region for Gp44 was cloned between NdeI and XhoI in pGAD GH to give pRJ3018, and full-length (FL) integrase or the CTD (residues 158 to 452) was cloned between NdeI and BamHI in pGBKT7 to give pRJ3017 or pRJ3014, respectively. Control plasmids expressing domains from yeast Sla1p and interacting (NPF) and noninteracting (NPA) versions of Kex2p were from Greg Payne (UCLA) (49). To assess relative growth on agar plates, freshly transformed PJ69-4A cells were grown overnight to stationary phase in synthetic complete (SC) medium without tryptophan and leucine, diluted to 1×10^7 cells/ml, and then 10-fold serially diluted. Dilutions were spotted onto SC-Trp/Leu plates with or without histidine and 3-amino-1,2,4-triazole and incubated at 30°C for 5 days.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/JB.00019-17.

SUPPLEMENTAL FILE 1, PDF file, 0.4 MB.

ACKNOWLEDGMENTS

We thank R. Calendar (University of California, Berkeley) for *Listeria* strains and DNA, G. Payne (University of California, Los Angeles) for yeast two-hybrid reagents, and R. Sharp (University of Pennsylvania) for assistance with protein expression and purification.

This work was supported by NIH grants GM038509 (to R.C.J.) and GM108751 (to G.D.V.D.).

REFERENCES

1. Smith MC. 2015. Phage-encoded serine integrases and other large serine recombinases. Microbiol Spectr 3:. https://doi.org/10.1128/microbiolspec.MDNA3-0059-2014.
2. Van Duyne GD, Rutherford K. 2013. Large serine recombinase domain structure and attachment site binding. Crit Rev Biochem Mol Biol 48:476–491. https://doi.org/10.3109/10409238.2013.831807.
3. Groth AC, Calos MP. 2004. Phage integrases: biology and applications. J Mol Biol 335:667–678. https://doi.org/10.1016/j.jmb.2003.09.082.
4. Smith MC, Thorpe HM. 2002. Diversity in the serine recombinases. Mol Microbiol 44:299–307. https://doi.org/10.1046/j.1365-2958.2002.02891.x.
5. Bibb LA, Hancox MI, Hatfull GF. 2005. Integration and excision by the large serine recombinase phiRv1 integrase. Mol Microbiol 55:1896–1910. https://doi.org/10.1111/j.1365-2958.2005.04517.x.
6. Breüner A, Brondsted L, Hammer K. 1999. Novel organization of genes involved in prophage excision identified in the temperate lactococcal bacteriophage TP901-1. J Bacteriol 181:7291–7297.
7. Ghosh P, Wasil LR, Hatfull GF. 2006. Control of phage Bxb1 excision by a novel recombination directionality factor. PLoS Biol 4:e186. https://doi.org/10.1371/journal.pbio.0040186.
8. Khaleel T, Younger E, McEwan AR, Varghese AS, Smith MC. 2011. A phage protein that binds phiC31 integrase to switch its directionality. Mol Microbiol 80:1450–1463. https://doi.org/10.1111/j.1365-2958.2011.07696.x.
9. Zhang L, Zhu B, Dai R, Zhao G, Ding X. 2013. Control of directionality in *Streptomyces* phage phiBT1 integrase-mediated site-specific recombination. PLoS One 8:e80434. https://doi.org/10.1371/journal.pone.0080434.
10. Lewis JA, Hatfull GF. 2003. Control of directionality in L5 integrase-mediated site-specific recombination. J Mol Biol 326:805–821. https://doi.org/10.1016/S0022-2836(02)01475-4.
11. Esposito D, Scocca JJ. 1997. Purification and characterization of HP1 Cox and definition of its role in controlling the direction of site-specific recombination. J Biol Chem 272:8660–8670. https://doi.org/10.1074/jbc.272.13.8660.
12. Mattis AN, Gumport RI, Gardner JF. 2008. Purification and characterization of bacteriophage P22 Xis protein. J Bacteriol 190:5781–5796. https://doi.org/10.1128/JB.00170-08.
13. Berntsson RP-A, Odegrip R, Sehlen W, Skaar K, Svensson LM, Massad T, Hogbom M, Haggard-Ljungquist E, Stenmark P. 2014. Structural insight into DNA binding and oligomerization of the multifunctional Cox protein of bacteriophage P2. Nucleic Acids Res 42:2725–2735. https://doi.org/10.1093/nar/gkt1119.
14. Abbani MA, Papagiannis CV, Sam MD, Cascio D, Johnson RC, Clubb RT. 2007. Structure of the cooperative Xis-DNA complex reveals a micronucleoprotein filament that regulates phage lambda intasome assembly. Proc Natl Acad Sci U S A 104:2109–2114. https://doi.org/10.1073/pnas.0607820104.
15. Xu Z, Thomas L, Davies B, Chalmers R, Smith M, Brown W. 2013. Accuracy and efficiency define Bxb1 integrase as the best of fifteen candidate serine recombinases for the integration of DNA into the human genome. BMC Biotechnol 13:87. https://doi.org/10.1186/1472-6750-13-87.
16. Xu Z, Brown WR. 2016. Comparison and optimization of ten phage encoded serine integrases for genome engineering in *Saccharomyces cerevisiae*. BMC Biotechnol 16:13. https://doi.org/10.1186/s12896-016-0241-5.
17. Brown WR, Lee NC, Xu Z, Smith MC. 2011. Serine recombinases as tools for genome engineering. Methods 53:372–379. https://doi.org/10.1016/j.ymeth.2010.12.031.
18. Keravala A, Groth AC, Jarrahian S, Thyagarajan B, Hoyt JJ, Kirby PJ, Calos MP. 2006. A diversity of serine phage integrases mediate site-specific recombination in mammalian cells. Mol Genet Genomics 276:135–146. https://doi.org/10.1007/s00438-006-0129-5.
19. Fogg PC, Colloms S, Rosser S, Stark M, Smith MC. 2014. New applications for phage integrases. J Mol Biol 426:2703–2716. https://doi.org/10.1016/j.jmb.2014.05.014.
20. Colloms SD, Merrick CA, Olorunniji FJ, Stark WM, Smith MC, Osbourn A, Keasling JD, Rosser SJ. 2014. Rapid metabolic pathway assembly and modification using serine integrase site-specific recombination. Nucleic Acids Res 42:e23. https://doi.org/10.1093/nar/gkt1101.
21. Bonnet J, Subsoontorn P, Endy D. 2012. Rewritable digital data storage in live cells via engineered control of recombination directionality. Proc Natl Acad Sci U S A 109:8884–8889. https://doi.org/10.1073/pnas.1202344109.
22. Siuti P, Yazbek J, Lu TK. 2013. Synthetic circuits integrating logic and memory in living cells. Nat Biotechnol 31:448–452. https://doi.org/10.1038/nbt.2510.
23. Mandali S, Dhar G, Avliyakulov NK, Haykinson MJ, Johnson RC. 2013. The site-specific integration reaction of *Listeria* phage A118 integrase, a serine recombinase. Mobile DNA 4:2. https://doi.org/10.1186/1759-8753-4-2.
24. Rutherford K, Yuan P, Perry K, Sharp R, Van Duyne GD. 2013. Attachment site recognition and regulation of directionality by the serine integrases. Nucleic Acids Res 41:8341–8356. https://doi.org/10.1093/nar/gkt580.
25. Loessner MJ, Inman RB, Lauer P, Calendar R. 2000. Complete nucleotide sequence, molecular analysis and genome structure of bacteriophage A118 of *Listeria monocytogenes*: implications for phage evolution. Mol Microbiol 35:324–340. https://doi.org/10.1046/j.1365-2958.2000.01720.x.
26. van Sinderen D, Luttinger A, Kong L, Dubnau D, Venema G, Hamoen L. 1995. *comK* encodes the competence transcription factor, the key regulatory protein for competence development in *Bacillus subtilis*. Mol Microbiol 15:455–462. https://doi.org/10.1111/j.1365-2958.1995.tb02259.x.
27. Rabinovich L, Sigal N, Borovok I, Nir-Paz R, Herskovits AA. 2012. Prophage excision activates *Listeria* competence genes that promote phagosomal escape and virulence. Cell 150:792–802. https://doi.org/10.1016/j.cell.2012.06.036.
28. Rutherford K, Van Duyne GD. 2014. The ins and outs of serine integrase site-specific recombination. Curr Opin Struct Biol 24:125–131. https://doi.org/10.1016/j.sbi.2014.01.003.
29. Yuan P, Gupta K, Van Duyne GD. 2008. Tetrameric structure of a serine integrase catalytic domain. Structure 16:1275–1286. https://doi.org/10.1016/j.str.2008.04.018.
30. Bai H, Sun M, Ghosh P, Hatfull GF, Grindley ND, Marko JF. 2011. Single-molecule analysis reveals the molecular bearing mechanism of DNA strand exchange by a serine recombinase. Proc Natl Acad Sci U S A 108:7419–7424. https://doi.org/10.1073/pnas.1018436108.
31. Olorunniji FJ, Buck DE, Colloms SD, McEwan AR, Smith MC, Stark WM, Rosser SJ. 2012. Gated rotation mechanism of site-specific recombination by phiC31 integrase. Proc Natl Acad Sci U S A 109:19661–19666. https://doi.org/10.1073/pnas.1210964109.
32. Ghosh P, Kim AI, Hatfull GF. 2003. The orientation of mycobacteriophage Bxb1 integration is solely dependent on the central dinucleotide of *attP* and *attB*. Mol Cell 12:1101–1111. https://doi.org/10.1016/S1097-2765(03)00444-1.
33. Smith MCA, Till R, Brady K, Soultanas P, Thorpe H, Smith MCM. 2004. Synapsis and DNA cleavage in phiC31 integrase-mediated site-specific recombination. Nucleic Acids Res 32:2607–2617. https://doi.org/10.1093/nar/gkh538.
34. Rowley PA, Smith MC, Younger E. 2008. A motif in the C-terminal domain of phiC31 integrase controls the directionality of recombination. Nucleic Acids Res 36:3879–3891. https://doi.org/10.1093/nar/gkn269.
35. Singh S, Ghosh P, Hatfull GF. 2013. Attachment site selection and identity in Bxb1 serine integrase-mediated site-specific recombination. PLoS Genet 9:e1003490. https://doi.org/10.1371/journal.pgen.1003490.

36. McEwan AR, Rowley PA, Smith MC. 2009. DNA binding and synapsis by the large C-terminal domain of phiC31 integrase. Nucleic Acids Res 37:4764–4773. https://doi.org/10.1093/nar/gkp485.
37. Pokhilko A, Zhao J, Ebenhoh O, Smith MC, Stark WM, Colloms SD. 2016. The mechanism of ϕC31 integrase directionality: experimental analysis and computational modelling. Nucleic Acids Res 44:7360–7372.
38. Ghosh P, Bibb LA, Hatfull GF. 2008. Two-step site selection for serine-integrase-mediated excision: DNA-directed integrase conformation and central dinucleotide proofreading. Proc Natl Acad Sci U S A 105:3238–3243. https://doi.org/10.1073/pnas.0711649105.
39. Fields S, Song O. 1989. A novel genetic system to detect protein-protein interactions. Nature 340:245–246. https://doi.org/10.1038/340245a0.
40. Savinov A, Pan J, Ghosh P, Hatfull GF. 2012. The Bxb1 gp47 recombination directionality factor is required not only for prophage excision, but also for phage DNA replication. Gene 495:42–48. https://doi.org/10.1016/j.gene.2011.12.003.
41. Saha S, Haggard-Ljungquist E, Nordstrom K. 1987. The Cox protein of bacteriophage P2 inhibits the formation of the repressor protein and autoregulates the early operon. EMBO J 6:3191–3199.
42. Esposito D, Wilson JC, Scocca JJ. 1997. Reciprocal regulation of the early promoter region of bacteriophage HP1 by the Cox and CI proteins. Virology 234:267–276. https://doi.org/10.1006/viro.1997.8646.
43. Sitlani A, Crothers DM. 1998. DNA-binding domains of Fos and Jun do not induce DNA curvature: an investigation with solution and gel methods. Proc Natl Acad Sci U S A 95:1404–1409. https://doi.org/10.1073/pnas.95.4.1404.
44. Wu HM, Crothers DM. 1984. The locus of sequence-directed and protein-induced DNA bending. Nature 308:509–513. https://doi.org/10.1038/308509a0.
45. Schuck P. 2000. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. Biophys J 78:1606–1619. https://doi.org/10.1016/S0006-3495(00)76713-0.
46. Vistica J, Dam J, Balbo A, Yikilmaz E, Mariuzza RA, Rouault TA, Schuck P. 2004. Sedimentation equilibrium analysis of protein interactions with global implicit mass conservation constraints and systematic noise decomposition. Anal Biochem 326:234–256. https://doi.org/10.1016/j.ab.2003.12.014.
47. Laue TM, Shah BD, Ridgeway TM, Pelletier SL (ed). 1992. Computer-aided interpretation of analytical sedimentation data for proteins. Royal Society of Chemistry, Cambridge, United Kingdom.
48. James P, Halladay J, Craig EA. 1996. Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. Genetics 144:1425–1436.
49. Howard JP, Hutton JL, Olson JM, Payne GS. 2002. Sla1p serves as the targeting signal recognition factor for $NPFX_{(1,2)}D$-mediated endocytosis. J Cell Biol 157:315–326. https://doi.org/10.1083/jcb.200110027.