Mutagenesis and Homology Modeling of the Tn*21* Integron Integrase IntI1[†]

Carolina Johansson,[‡,§] Lars Boukharta,[§,||] Jens Eriksson,[‡] Johan Åqvist,[||] and Lars Sundström*[,‡]

Department of Medical Biochemistry and Microbiology, Uppsala University, BMC, Box 582, 751 23 Uppsala, Sweden, and Department of Cell and Molecular Biology, Uppsala University, BMC, Box 596, 751 24 Uppsala, Sweden

Received October 30, 2008; Revised Manuscript Received January 8, 2009

ABSTRACT: Horizontal DNA transfer between bacteria is widespread and a major cause of antibiotic resistance. For logistic reasons, single or combined genes are shuttled between vectors such as plasmids and bacterial chromosomes. Special elements termed integrons operate in such shuttling and are therefore vital for horizontal gene transfer. Shorter elements carrying genes, cassettes, are integrated in the integrons, or excised from them, by virtue of a recombination site, *attC*, positioned in the 3′ end of each unit. It is a remarkable and possibly restricting elementary feature of *attC* that it must be single-stranded while the partner target site, *attI,* may be double-stranded. The integron integrases belong to the tyrosine recombinase family, and this work reports mutations of the integrase IntI1 from transposon Tn*21*, chosen within a well-conserved region characteristic of the integron integrases. The mutated proteins were tested for binding to a bottom strand of an *attC* substrate, by using an electrophoresis mobility shift assay. To aid in interpreting the results, a homology model was constructed on the basis of the crystal structure of integron integrase VchIntIA from *Vibrio cholerae* bound to its cognate *attC* substrate VCRbs. The local stability and hydrogen bonding network of key domains of the modeled structure were further examined using molecular dynamics simulations. The homology model allowed us to interpret the roles of several amino acid residues, four of which were clearly binding assay responsive upon mutagenesis. Notably, we also observed features indicating that IntI1 may be more prone to base-specific contacts with VCRbs than VchIntIA.

Integrons are widespread recombination elements with a capacity for gene capture through site-specific recombination (*1*). The elements were originally found on plasmids in human pathogens where they carry a broad variety of different antibiotic resistance genes (*2*−*5*). More recent data suggest that resistance integrons are derived from an ancestral pool of alleles in the chromosomes of certain bacteria, for instance *Vibrio cholerae* (*6*). Such chromosomal integrons, called superintegrons, may contain hundreds of gene cassettes of diverse or unknown functions. Unlike integrative conjugative elements (ICE) (*7*), integrons themselves are formally immobile, but transposons that are typically found on large conjugative plasmids or genetic islands make integrons mobilizable and connected to horizontal transfer (*8*). With their immense and varied repertoire of movable cassettes, integrons comprise a strong resource in evolution and diversification of resistance.

Integration and excision of gene cassettes (*9*) are mediated by an integron-encoded site-specific recombinase, IntI, of the tyrosine family. This site-specific recombination occurs between two sites of either *attI* or *attC* classes (*10*−*12*) (Figure 1). Site-specific recombination is a nonreplicative and phosphate energy-conservative process for reciprocal DNA cleavage and strand exchange between defined sites (*13*). It is widely used in prokaryotes and their viruses, for a variety of biological purposes such as phage lysogenization, replicon stabilization, and expression phase switching (*14*). Site-specific recombination mediated by the tyrosine recombinases follows a two-step cleavage and rejoining pathway in which the enzyme is bound directly to the DNA in the intermediate. The first step of this transesterification utilizes a tyrosine side chain oxygen as the nucleophilic agent. In

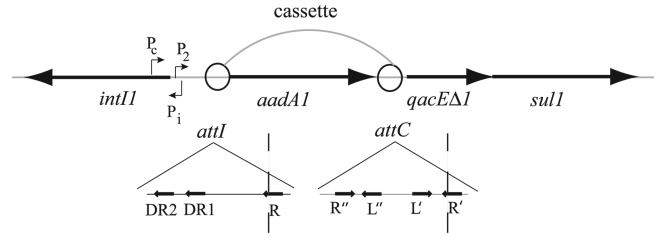

FIGURE 1: Schematic picture of the organization of the integron in Tn*21* (GenBank entry X12870), carrying the *aadA1* cassette encoding streptomycin/spectinomycin resistance and the following *qacEΔ1-sul1* combination downstream of *attC* (*5*). The integron integrase is encoded in the 5′-conserved sequence, to the left. Genes are shown in boldface horizontal arrows and recombination sites as empty circles. The curve shows the extent of the *aadA1* cassette. P_c and P_2 are promoters used for expression of integrated gene cassettes and P_i for integrase. In the bottom panel, the two primary recombination sites, *attI* and *attC*, are enlarged for visualization of their different repeat organization. In the *attI* site, two direct repeats (DR2 and DR1) are found upstream of repeat R that includes the recombination crossover point. The *attC* site on the other hand is almost a perfect palindrome organized in two subsites of dyad symmetry (repeats R″ and L″ and repeats L′ and R′). Vertical dashed lines mark cassette borders.

[†] This work was supported by the EU/FP6-funded DRESP2 project and AFA Health Fund (to L.S.), the Swedish Research Council (VR) (to L.S. and J.Å.), and the Anna-Maria Lundins Fund (to C.J.).

* To whom correspondence should be addressed. Phone: +46 18 471 4115. Fax: +46 18 471 4673. E-mail: Lars.Sundstrom@imbim.uu.se.

[‡] Department of Medical Biochemistry and Microbiology.

[§] These authors contributed equally to this work.

[||] Department of Cell and Molecular Biology.

10.1021/bi8020235 CCC: $40.75
Published on Web 02/06/2009

the following step, the reactive phosphotyrosyl link in the DNA−protein conjugate is attacked by the 5′-OH group on the phosphate-free DNA end in the partner site. Two pairs of such reactive intermediates are needed for completing the recombination progressing via a Holliday junction intermediate. Similar phosphoryl transfer mechanisms are used by topoisomerases, relaxases, and proteins involved in telomere maintenance in instances where chromosomes have a linear form (*15, 16*).

Tyrosine recombinases contain structurally conserved parts (see below) but, nevertheless, differ extensively in amino acid sequence and in their requirement for accessory proteins and recognition sites. Biochemical studies of several of the tyrosine recombinases have shown a functional separation into evolutionary distinctive N-terminal and C-terminal domains. The C-terminal domain contains the most conserved amino acids and recognizable sequence motifs that are the primary hallmarks of the tyrosine recombinase family (*17*−*19*). Prominent features are the six noncontiguous but highly conserved amino acids (RKHRHY) that form the catalytic site within the C-terminal domain (*17, 20, 21*). Structural studies have shown that the fold of the entire C-terminal domain is well-conserved even with a low level of sequence identity outside the active site (*22*).

Several structures of complete synaptic complexes, monomeric DNA complexes, and protein in the absence of DNA are now available (e.g., HP1, λ Int, Cre, XerD, Flp, VchIntIA, human topoisomerase I, and vaccinia virus topoisomerase) (*23*−*30*). Accumulating data have suggested that all tyrosine recombinases use minor variations of the same basic mechanism. The integron integrases, containing the most conserved sequence motifs, have therefore been expected to act like the other tyrosine recombinases. Interestingly, in an alignment of all tyrosine recombinases, the integron integrases differ from the other sequences by a large polypeptide insertion. This insertion is located between helices αI and αJ and, more specifically, between relatively well conserved patches II and III. This anomaly was first described by Nunes-Düby et al. (*18*), and its possible function was later emphasized (*31*). Recently, the crystal structure of the integron integrase VchIntIA from *V. cholerae* bound to an oligo reconstruction of its *attC* substrate (VCR_{bs})[1] has been reported (*32*). Changes in the unique amino acid insertion have been shown to affect both the binding (*in vitro*) of *attI* and recombination (*in vivo*) (*31*) (Table 1). According to the VchIntI1A−VCR_{bs} structure, there is one extra short α-helix within this essential region named $αI_2$ with an important role in synapse formation.

Unlike several other family members, IntIs can mediate the exchange of DNA between two architecturally distinct sites, *attI* and *attC* (Figure 1). It has been demonstrated that integron integrases bind to the bottom strand of an *attC* site folded into a hairpin with bulges important for recognition (*32*−*34*). The importance of secondary structure features rather than sequence for recognition explains how extensively varied *attC* sites are cross-recognized by different integron integrases (*5, 6, 21, 35, 36*). In the crystal structure of the *V. cholerae* integron integrase, four VchIntIA monomers are bound to two antiparallel VCR_{bs} molecules that are bulged duplex reconstructions of a folded bottom strand of an *attC* substrate for VchIntIA (*32*). The currently prevalent recombination model suggests a single-stranded recombination pathway that stops at the Holliday junction intermediate, leaving it to be resolved to product by yet unknown cellular mechanisms (*34*).

Table 1: List of Previously Used Mutants of IntI1 and Their Effect on DNA Binding to *attI* and Recombination (*31, 55*)[a]

IntI1 mutation	DNA binding	recombination
E121D	++	++
E121K	+	+
R146K	++	−
R146E	−	−
R146I	−	−
R146V	−	−
K171E	++	+
K171I	++	+
K171Q	+	+
K171R	++	+
K171V	++	+
ΔALER215	++	−
K219E	+++	−
K219I	++	++
K219W	+++	−
F233L	+	+
F233R	+	+
F233Y	++	+
W229G	−	−
W229R	+	−
H277D	−	−
H277L	++	−
H277R	+++	−
H277Y	+++	+
R280G	−	−
R280E	−	−
R280K	++	+
G302A	++	+
G302R	++	−
Y312S	++	−
Y312F	++	−

[a] Key: −, negative; +, weaker than the wild-type protein; ++, wild-type efficiency; +++, stronger than the wild-type protein.

In this work, we have modified the type 1 integron integrase IntI1 from Tn*21* into 20 protein variants and tested them for binding to a bottom strand of an *attC* substrate, named $attC_{bs}$ (see Materials and Methods), by using an electrophoresis mobility shift assay (EMSA) (Figure 2). To rationalize these results, we have also attempted structural comparison with VchIntIA, which is 45% identical to IntI1 (Figure 2), using homology modeling based on the published structure of VchIntIA.

MATERIALS AND METHODS

Mutagenesis. Plasmids encoding the various mutations in His-tagged IntI1 were constructed by using the Stratagene QuickChange site-directed mutagenesis kit (AH diagnostics). Thirty nanograms of plasmid p2352 (*33*) containing the gene for IntI1 of Tn*21* fused to the His·Tag sequence in a pET19b vector was used as the template. Two complementary primer pairs containing the desired mutations, designed according to the manufacturer's instructions, were used to construct each gene mutant. The complete sequencing of all mutant genes revealed no discrepancy, except for the desired mutations, when compared to the wild-type integrase gene.

Crude Extraction Procedure. The plasmids were transformed into the *Escherichia coli* production strain BL21(DE3)-

[1] Abbreviations: IntI, integron integrase; VchIntIA, *V. cholerae* integron integrase; VCR_{bs}, *V. cholerae* repeat bottom strand; EMSA, electrophoresis mobility shift assay; MD, molecular dynamics.

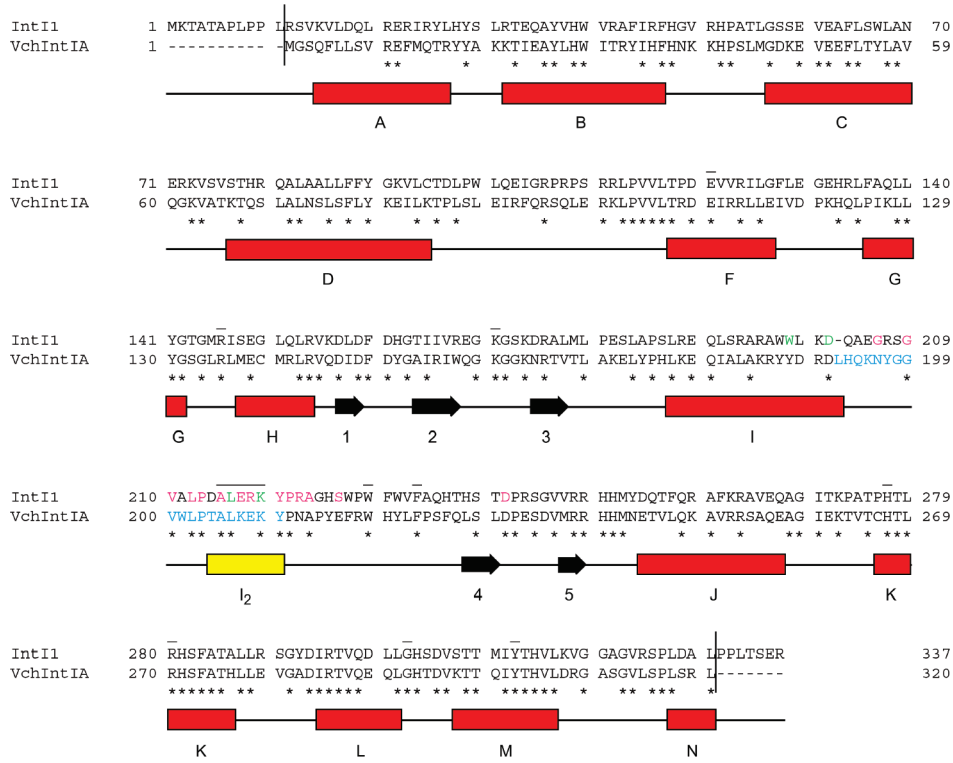

FIGURE 2: Sequence alignment of IntI1 (GenBank entry CAA31361) and VchIntIA (GenBank entry AAC38424). The two proteins are 45% identical in sequence. The vertical lines mark the start and stop of the amino acid sequence of IntI1 used for homology modeling. Secondary structure elements for VchIntIA earlier presented by MacDonald et al. (*32*) are shown, and α-helix I_2 found in the unique insertion sequence is highlighted in yellow. The cyan stretch of amino acids in VchIntIA indicates the unique insertion region. The mutations that were made in IntI1 in this study are marked in green or magenta. The color indicates if the mutant IntI1 containing the substitution was responsive (green) or nonresponsive (magenta) in the EMSA. Amino acids that previously have been mutated (see Table 1) are overlined.

plysS. To perform crude extract analysis, a loop of cells, from the production strain, was grown overnight in a 5 mL culture of LB medium, containing 100 μg/mL ampicillin and 170 μg/mL chloramphenicol. One milliliter of overnight culture was reinoculated in 100 mL of LB medium, containing the same antibiotic concentrations listed above, and was allowed to grow at 37 °C to mid-late exponential phase (OD_{600} = 0.69) and then cooled to 16 °C on ice. T7 polymerase-mediated integrase expression was induced by addition of isopropyl β-D-thiogalactopyranoside (IPTG) at 1 mM. After induction, retarded growth was continued for 20 h at 16 °C with gentle shaking. The cells were then harvested by centrifugation at 5000 rpm for 5 min in 4 °C. Pellets were frozen at −20 °C and stored for no longer than a week. The frozen pellets were resuspended on ice in 2.25 mL of phosphate buffer [20 mM Na_2HPO_4-NaH_2PO_4 (pH 7.2), 200 mM NaCl, and 1 mM imidazole] supplemented with 250 μL of 1% Triton X-100 and 25 μL of a 100 mg/mL solution of the protease inhibitor Pefablock SC (AEBSF) (Roche Diagnostics). The cells were sonicated on ice for 4 × 30 s with a 30 s cooling period between the pulses. Insoluble material was separated by centrifugation at 15000 rpm for 30 min at 4 °C. The supernatant was immediately supplemented with 11% of 87% glycerol (MERCK) and stored at −70 °C. The presence of integrase in the crude extracts was analyzed via SDS−PAGE. All the crude extracts exhibited similar amounts of the 38 kDa integrase band, the only exception being the negative controls.

Labeling of attC Oligos and Assaying for IntI1 Binding. An oligonucleotide, named $attC_{bs}$, representing the integrase binding bottom strand of the *attC* site of the *aadA1*−*qacE* cassette junction in Tn*21* was used as a substrate and labeled at its 5′-terminus with radioactive phosphate transferred from [γ-^{32}P]ATP (Amersham) by polynucleotide kinase. IntI1−DNA complexes were assayed by reduction of their electrophoretic mobilities on a 6% nondenaturing polyacrylamide gel [AccuGel 29:1 acrylamide:bisacrylamide (In Vitro)]. Both purified integrase [purification described by Johansson et al. (*33*)] and the crude extract of unmutated wild-type integrase were used as positive controls. Labeled oligo incubated in the absence of the crude extract and in the presence of crude extracts of BL*21*(DE3)plysS with and without the pET19b vector served as negative controls. First, all of the mutated integrases and controls were examined on separate gels with a wide range of crude extract amounts used (from 0.5 to 10 μL) to determine the optimal amount for comparison. Individual reactions were set up as follows: 1 μL (8 nM, ∼200−500 cps) of labeled oligo was incubated with various amounts of crude extract supplemented with phosphate buffer to a volume of 5 μL (for stringent NaCl concentrations). The 20 μL binding mixtures were supplemented additionally with 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 5% glycerol, 0.20 mM EDTA, 1 mM CHAPS, 1 mM DTT, 0.7 μg of BSA, and 1 μg of poly(dI-dC)·poly(dI-dC). Incubation for binding took place for 15 min at 30 °C. Following incubation, 2 μL of loading dye was added [0.2% xylene cyanol, 0.2% bromophenol blue, 50% saccharose, and 100 mM EDTA (pH

7.9)], and samples were loaded on the polyacrylamide gel and run at 4 °C in a 1× TBE buffer.

An EMSA under denaturing conditions was performed as described above but with the following changes. The reactions were stopped by addition of 2 μL of stop solution [50 mM Tris-HCl (pH 7.5) and 1% SDS] before the samples were loaded on a 6% polyacrylamide gel supplemented with 0.1% SDS. The running buffer contained 1× TBE and 0.1% SDS. The gel was run at room temperature.

attC$_{bs}$ used in our binding assays is not identical to VCR$_{bs}$ found in the crystal structure and hence also in our homology model. However, the sites differ at only eight of 27 positions in the four binding repeats (R″, L″, L′, and R′), and those were all nonconserved according to Biskri et al. (*36*).

Alignment, Homology Modeling, and Molecular Dynamics Simulations. The amino acid sequence of IntI1 was taken from GenBank entry X12870 and the VchIntIA−VCR$_{bs}$ sequence with a 2.8 Å resolution structure from Protein Data Bank entry 2A3V (*32*). These two sequences were aligned with T-Coffee (*37*). IntI1 and VchIntIA share 45% amino acid sequence identity, which should be enough for a reliable model. We have used subunits B and C (nonattacking and attacking, respectively) as templates and mainly looked at those parts that are close to the bound DNA.

SWISS-MODEL (*38*) was then used for straightforward homology modeling. The program uses a combination of the provided template structure, constraint programming, scoring functions, and fragment libraries to predict the position of all the residues in the homology model. The aligned sequences were submitted in the alignment mode. After manual adjustments in subunits B and C of IntI1, we aligned them with subunits D and A from *V. cholerae* for assembly to a full tetrameric complex.

With the high degree of indentity between the two proteins, the overall fold is expected to be the same and molecular dynamics (MD) simulations were used to examine the local stability and hydrogen bonding network of the modeled structure for some key regions. All the simulations were carried out with Q (*39*), using the Charmm22 force field (*40*). Five different spherical simulation systems were used. Four had a radius of 20 Å and were centered on the catalytic tyrosine, the bulges, and the center of the protein−DNA interface of subunit C. The fifth system had a radius of 30 Å, and was centered on the protein−DNA interface of subunit B. The systems were solvated with TIP3P waters (*41*), and the water surfaces of the spheres were subjected to radial and polarization restraints (*42*), to mimic bulk water at the sphere boundary. Nonbonded interaction energies were calculated up to a 10 Å cutoff, while long-range electrostatics were treated using a multipole expansion (*43*). Atoms outside the simulation spheres were restrained to their initial positions and only interacted with the system through bonds, angles, and torsions. All systems were heated from 1 to 310 K using a stepwise scheme. Positional restraints on the heavy atoms were then gradually released from 25 to 0 kcal mol^{-1} Å^{-2}, and the temperature relaxation time was changed from 0.2 to 100 fs during this heating process. The production phase consisted of 300 ps following an equilibration period of 10 ps. Equilibration and production were carried out at 310 K, using a time step of 1 fs. Structural averages of the molecular trajectories were generated and examined along with the dynamic behavior of the systems.

Table 2: Results of an EMSA and a Gel Retardation Assay Performed under Denaturing Conditions for Detect the Formation of a Covalent Linkage between *attC*$_{bs}$ and Integron Integrase[a]

	DNA binding	covalent linkage
wild type	++	++
W199A	−	−
W199Y	++	++
D202A	±	±
G206D	++	not tested
G209D	++	not tested
V210T	++	not tested
L212F	++	not tested
P213R	++	not tested
A215E	++	not tested
L216R	+	+
E217A	++	not tested
R218A	++	not tested
K219W	±	±
Y220S	++	++
Y220F	++	not tested
P221A	++	not tested
R222A	++	not tested
A223V	++	not tested
S226A	++	not tested
D214G	++	not tested

[a] Key: −, negative; ±, strongly reduced compared to the wild-type protein; +, weaker than the wild-type protein; ++, wild-type efficiency. For more details, see Figure 3A,B.

RESULTS

Mutagenesis Strategy. A region of conserved amino acids encompasses an insertion with the ALER motif previously highlighted by Messier and Roy (*31*) for being a characteristic of all integron integrases (Figure 2). In this work, a generous 65-amino acid segment covering positions 187−251 in IntI1 from Tn*21* was aligned with a panel of integron integrases, and eighteen residues were chosen for directed mutagenesis to add information to those mutations previously described in the literature (Table 1). The 18 altered amino acids were W199, D202, G206, G209, V210, L212, P213, A215, L216, E217, R218, K219, Y220, P221, R222, A223, S226, and D241.

Responsive Mutations in an EMSA. Protein extracts of the mutated IntI1 variants were screened against positive and negative controls for evaluating instances in which there was a reduced level of binding to the *attC* site. All 20 mutations at the 18 amino acids were tested for the mobility shift of a ^{32}P-labeled DNA oligonucleotide representing the bottom strand of the only *attC* site of Tn*21*. The presence of the pET19b vector alone yielded an extract that shifted the DNA into a background pattern of unspecific affinities for the mobility-shifted DNA fragment. However, the specific band shifts due to integrase binding were not difficult to distinguish from unspecific bands. The most obvious effect by any single mutation was the lost band shift with IntI1W199A. Further, very faint shifted bands were seen with D202A and K219W, and a small amount of shifted DNA remained with L216R. In summary, the EMSA results suggest that all four mutations to varying degrees reduced the level of binding of the protein to the bottom strand of *attC* (*attC*$_{bs}$) as compared to binding of wild-type IntI1 (Table 2 and Figure 3A). At the bottom of the gel there was some unwanted DNA degradation most likely due to the presence of nuclease activity in the extract. However, it was estimated to have no influence on conclusions drawn from the experiment.

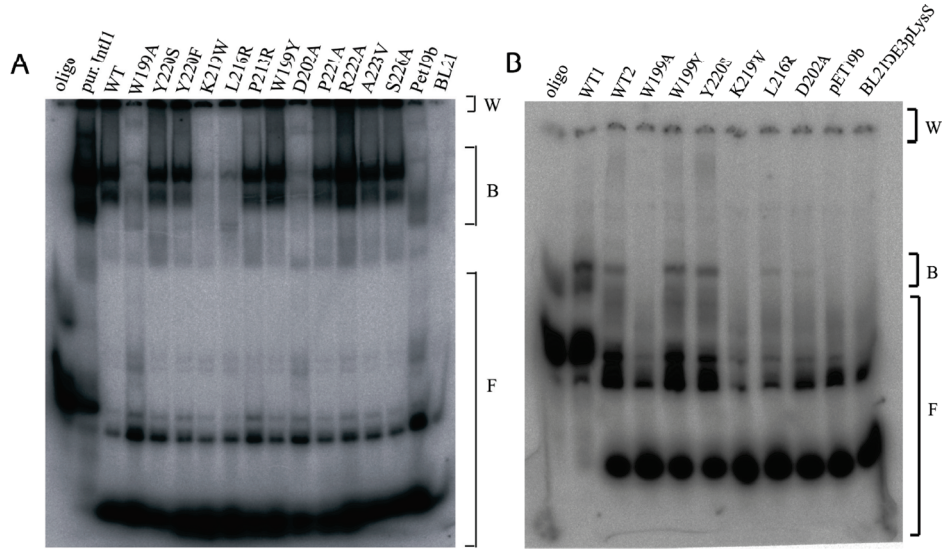

FIGURE 3: (A) Autoradiogram of EMSA results. Binding of purified and crude extract of wild type (WT) IntI1 to $attC_{bs}$ was compared to binding of 12 of the 20 mutated proteins to the same substrate. IntI1W199A, IntI1D202A, IntI1L216R, and IntI1K219W were clearly responsive in that binding was lost or greatly weakened. Those mutations that are not shown in the autoradiogram were nonresponsive; 1 μL of each crude extract was mixed with 1 μL of oligo (~8 nM). The amount of purified IntI1 in lane 2 corresponds to approximately 3 μM. Crude extracts of BL21DE3pLysS with and without pET19b mixed with oligo were used as controls. The two lowermost bands in the lanes that contain crude extracts are oligo degraded by nucleases present in the extracts. For more details, see Materials and Methods. (B) Gel retardation assay performed under denaturing conditions to detect the formation of a covalent linkage between $attC_{bs}$ and integron integrase. Oligo (1 μL, ~8 nM) was mixed with 1 μL of crude extract. WT1 = 5 μL of purified wild-type IntI1 corresponding to ~3 μM; WT2 = 1 μL of crude extract of wild-type IntI1. No band was observed with IntI1W199A indicating that no covalent linkage had been formed. Weak bands were observed with IntI1D202A, IntI1L216R, and IntI1K219W. The two lowermost bands in the lanes that contain crude extracts are oligo degraded by nucleases present in the extracts. The results are also presented in Table 2. Abbreviations: B, protein−DNA complexes; F, free and degraded single-stranded DNA; W, aggregated DNA that remained in the wells.

Nonresponsive Mutations in an EMSA. The majority of 16 tested mutations in IntI1 did not diverge from the EMSA response of the wild-type integrase, indicating that they bound equally well to $attC_{bs}$ like wild-type IntI1. They make up a list of regional amino acids of low importance for binding of the protein to the bottom strand *attC* DNA oligomer: W199Y, G206D, G209D, V210T, L212F, P213R, A215E, E217A, R218A, Y220S, Y220F, P221A, R222A, A223V, S226A, and D241G (Table 2 and Figure 3A). It was noticed that three of these nonresponsive mutations, A215E, E217A, and R218A, are in fact part of the ALER motif previously described and examined by Messier and Roy (*31*) (Table 1). Interestingly, these authors reported that deletion of the ALER motif did not affect binding to the double-stranded site, *attI,* an efficient partner site for recombination with the hairpin-recognized *attC* site studied here, while the deletion at the same time caused a deficiency in recombination. The deficiency in recombination might be due to the mutation in ALER at L216, one of the residues that according to our data indeed is involved in binding to $attC_{bs}$ (see above).

Catalytic Activity of Some Mutants Tested by SDS−PAGE. The catalytic activities of the altered IntI1 proteins with the four mutations causing perturbed protein binding to the bottom strand *attC* DNA (W199A, D202A, L216R, and K219W) were tested by using an EMSA also under denaturing conditions. Under denaturing SDS conditions, a shifted band is indicative of a covalent phosphotyrosine linkage between DNA and protein (*33*). W199Y does not alter the DNA binding of the wild-type protein that is entirely lost by mutagenesis of the same amino acid to alanine (W199A). Like another binding-neutral mutation, Y220S, it allowed the protein−DNA complexes to be formed and bands shifted as a normal wild-type IntI1-$attC_{bs}$ under both gel conditions, indicating that these mutations affected neither binding nor the catalytic activity. However, there was no band shift in the denaturing gel with the W199A mutation in IntI1 indicative of the lack of phosphotyrosyl linkage with the DNA. The observations with W199Y and W199A are consistent with the unidirectional relation between binding and catalysis that was reported by Johansson et al. (*33*). Reduced catalysis in step with loss of DNA binding was observed with the mutations D202A, L216R, and K219W. In the presence of either of these mutations, weakened shifted bands remained under denaturing as well as nondenaturing electrophoresis conditions (Table 2 and Figure 3B).

Homology Modeling of IntI1 against the VchIntIA−VCR_{bs} Synaptic Complex. In the crystal structure used as a template, two of the four VchIntIA molecules bound to two antiparallel VCR_{bs} duplexes are in the active conformation (able to attack DNA) while the other two are nonattacking subunits. This division of labor is determined by the site organization in which the positions of two extrahelical bases, G20″ and T12″, play an important role in the correct assembly of the monomers. The protein−DNA interactions are otherwise mainly nonspecific (*32*).

Due to the high degree of sequence indentity, our homology model of IntI1 has the same fold as VchIntIA (Figure 4A,B), and we have used MD simulations for local relaxation of the model to examine structural features such as hydrogen bonding. To begin, we looked at the six highly conserved amino acids among tyrosine recombinases that in IntI1 correspond to R146 (R135 in VchIntIA), K171 (K160), H277 (H267), R280 (R270), H303 (H293), and the catalytic Y312 (Y302) (Figure 2). All these amino acids aligned very well with the existing structure, both in the two attacking subunits and in the two nonattacking subunits. The active

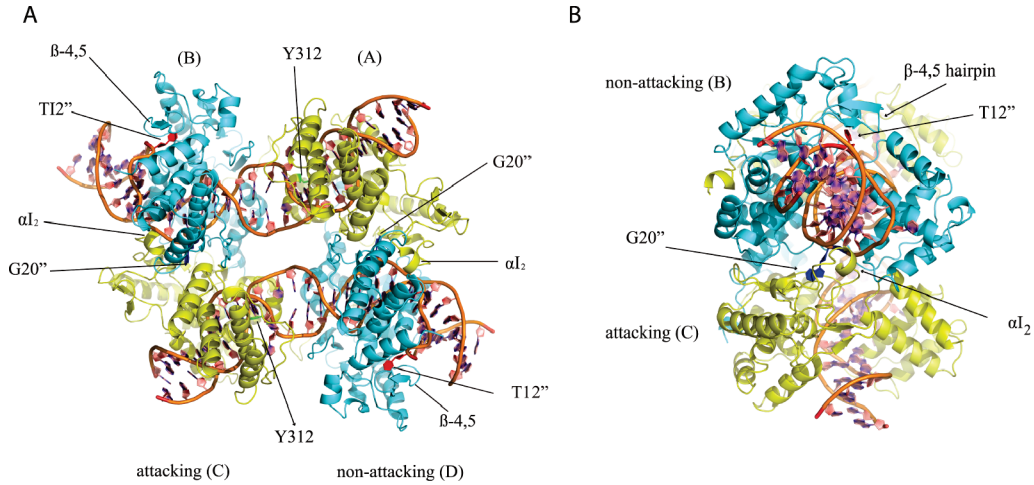

FIGURE 4: Architecture of our homology model of four IntI1 monomers from Tn*21* bound to two antiparallel VCR_{bs} duplexes using the VchIntIA−VCRbs structure as a template (*32*): (A) N-terminal view and (B) orthogonal view. The attacking subunits are colored yellow and the nonattacking subunits cyan. VCR_{bs} is colored pink. Extrahelical bases T12″ and G20″ are colored red and blue, respectively, and the attacking Y312 is colored green. α-Helix I_2 and the *β*-4,5 hairpin are pointed out in attacking and nonattacking subunits, respectively.

site is stable and the hydrogen bonding network conserved during the MD simulation of the active site in the IntI1 model. The amino acids in the active site are all derived from the same subunit, which is provided in *cis*. The attacking subunits make important DNA contacts also in *trans* with the extrahelical base G20″ through interactions with Q145, W157, K209, Y210, and W219 in VchIntIA that in IntI1 correspond to K156, R168, K219, Y220, and W229, respectively (Figure 5A). Q145/K156 and W157/R168 contain different amino acids at equivalent positions in the two proteins but seem to fit almost equally well in the structure model, which is also supported by the MD simulations. Residues 204−210 in VchIntIA (TALKEKY) correspond to residues 214−220 (DALERKY) in IntI1 and constitute the αI_2 helix situated in the unique insertion region. This helix was shown by MacDonald et al. (*32*) to be important for synapse formation and contains two of the important amino acids that interact with G20″ listed above. The amino acid differences of T versus D, K versus E, and E versus R should not disturb the formation of the helix. The extrahelical base T12″ is stabilized by *cis* interactions within the two inactive subunits and is important for DNA site recognition. The *β*-4,5 hairpin was shown earlier to interact with T12″ in a way that T12″ becomes inserted between two stacked histidines (H240 in VchIntIA; H250 in IntI1 and H241; H251) at one end and a highly conserved proline (P232; P242) at the other end, to form a tight nonpolar nucleotide−protein interface (*32*) (Figure 5B). The amino acids mentioned above are highly conserved in the two proteins, but a few residues found downstream of H250, H251, and P242 in IntI1 differ from VchIntIA, and these needed manual adjustments to fit into the structure. The binding between IntI1 and the T12″ bulge is stable during molecular dynamics.

Rationalization of Responsive and Nonresponsive Mutations in IntI1. The bottom strand of the *attC* site used in our EMSA study originates from the *aadA1*−*qacE* cassette junction in Tn*21*. It can be folded into a structure similar to that of VCR_{bs}; however, it does not contain the extrahelical base T16″ between T12″ and G20″ that is present in the latter (Figure 5C). This supplementary bulging nucleotide is conserved among VCRs, but not between *attC* sites in general (Supplemental Figure 1 in ref *34*). In the VchIntIA−VCR_{bs} crystal structure and the IntI1−VCR_{bs} homology model, T16″ is situated at the surface of nonattacking subunit D/B. A molecular dynamics simulation was carried out to investigate how deletion of T16″ would affect the local DNA and protein structure. The average structure of the simulation showed that the IntI1 conformation and binding to the T12″ and G20″ extrahelical bases were conserved regardless of the T16″ bulge. This result is also in agreement with the assumption that homologous integrases will utilize the same conserved binding pockets to bind structural features that are conserved among *attC*s.

The amino acid W199 was mutated into an alanine, which is aliphatic, small, and uncharged, to eliminate possible stacking and hydrophobic interactions. The mutation IntI1W-199A was responsive in the EMSA. When it was mutated back into an aromatic residue (W199Y), it regained its affinity for $attC_{bs}$ (Table 2 and Figure 3A,B). Aromaticity or a bulky, ringlike structure of the amino acid residue at position 199 seems to be important. In alignments of integron integrases, either a tryptophan, histidine, tyrosine, or phenylalanine is commonly found in this position. In our homology model, W199 is involved in hydrophobic packing between α-helix I and the loop connecting αI_2 and the *β*-4,5 hairpin (Figure 5C). It is situated close to the protein surface and far from the bound DNA. The responsiveness could be explained by structural changes of the αI_2−*β*-4,5 hairpin region, and W199 could be important for correct folding.

Position 202 is conserved among integron integrases. By the mutation D202A, the amino acid is changed from a negative to an aliphatic residue that in our model is positioned close to the surface of the protein. This residue has no contact with DNA or with any of the other subunits. D202 is buried between αI and the *β*-4,5 hairpin and coordinates a salt bridge to R249 and is near R207, and with its negative charge, it could be expected to have a role in stabilizing the conformation of the protein (Figure 5D).

L216 is rather conserved among integron integrases. L216 was substituted with a positively charged and quite large arginine residue. L216R was responsive but still retained

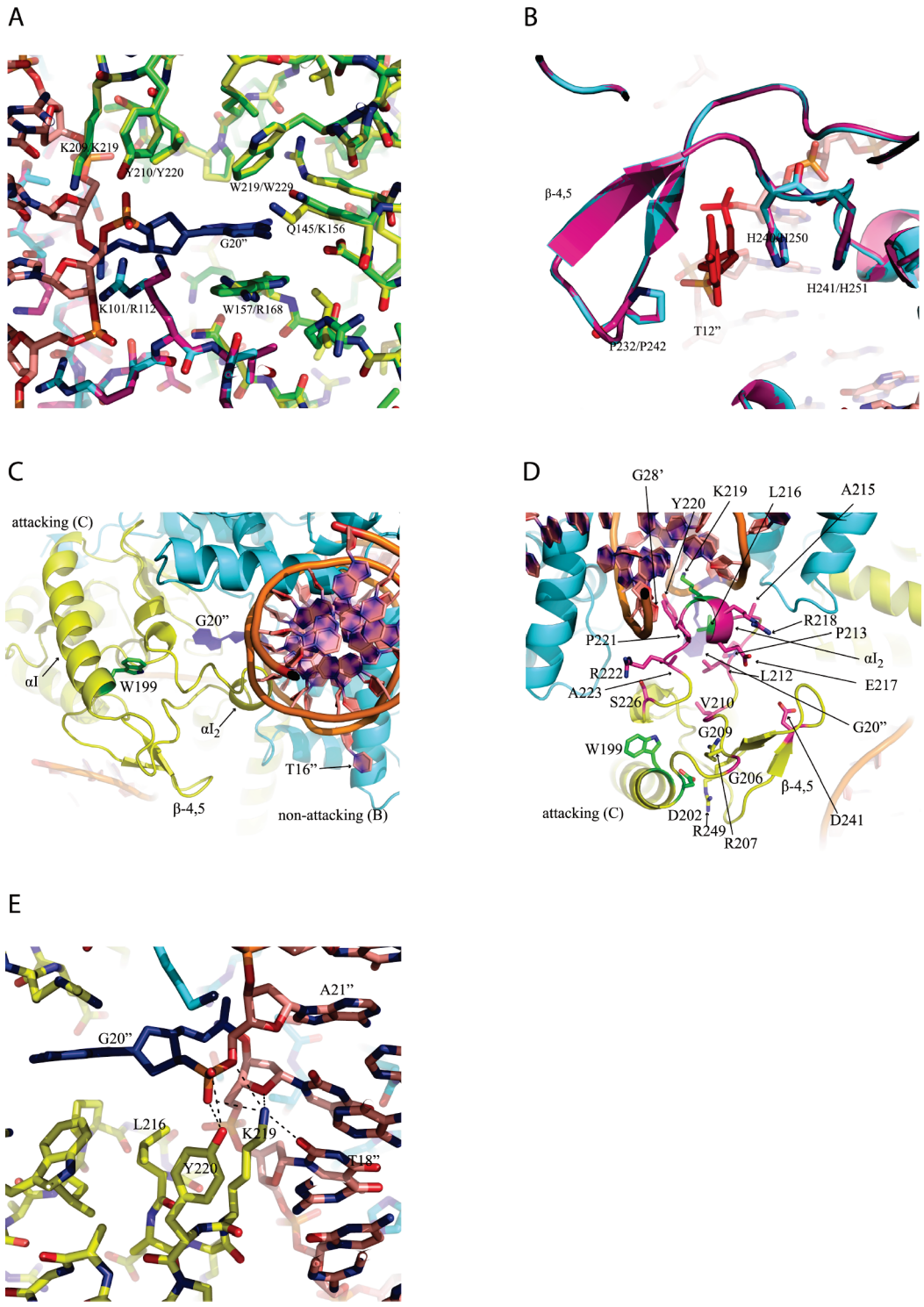

FIGURE 5: (A) View of the modeled IntI1 superimposed on the VchIntIA−VCR_{bs} crystal structure by MacDonald et al. (*32*) showing the *trans* interactions with the extrahelical base G20″ (blue). DNA is colored pink at the left, the attacking subunit C of VchIntIA green, the attacking subunit C in the model of IntI1 yellow, the nonattacking subunit B of VchIntIA magenta, and the nonattacking subunit B in the model of IntI1 cyan. Involved amino acids are highlighted in VchIntIA and IntI1. (B) View of the modeled IntI1 superimposed on the VchIntIA−VCR_{bs} crystal structure by MacDonald et al. (*32*) showing the *cis* interactions with the extrahelical base T12″ (red). DNA is colored pink, the nonattacking subunit B of VchIntIA magenta, and the nonattacking subunit B in the model of IntI1 cyan. Involved amino acids are highlighted in VchIntIA and IntI1. (C) Ribbon diagram from the IntI1 homology model showing the role of W199 (green) in the hydrophobic packing between α-helix I and the loop connecting α-helix I_2 and the *β*-4,5 hairpin. The nonattacking subunit is colored cyan, the attacking subunit yellow, and G20″ blue and DNA at the right. (D) Ribbon diagram from the IntI1 homology model showing the amino acids that have been mutated in this work colored magenta and green. The amino acids colored magenta (G206, G209, V210, L212, P213, A215, E217, R218, Y220, P221, R222, A223, S226, and D241) were found to be nonresponsive in the EMSA, and those colored green (W199, D202, L216, and K219) were found to be responsive. Residues R249 and R207 are colored yellow. (E) Detailed view of the *trans* interactions in the IntI1 homology model made by L216, K219, and Y220 with T18″, G20″ (in blue), and A21″ in DNA at the right.

some binding to $attC_{bs}$. This residue is found in αI_2 and is part of the ALER motif deleted by Messier and Roy (*31*). Deletion of this motif resulted in retained *attI* binding, but recombinational deficiency. In our model, L216 forms a

hydrophobic interaction with G20″. A substitution of the leucine with the larger and positively charged arginine residue affects the surroundings of G20″ (Figure 5E).

K219 is conserved among integron integrases, found in α-helix I_2, and adjacent to universally conserved Y220. K219 (K209 in VchIntIA) has been shown in the VchIntIA−VCR_{bs} crystal structure to hydrogen bond to DNA in the vicinity of the bulge G20″. The position was changed from a positively charged residue to a bulky aromatic tryptophan which resulted in loss of binding. In our model, it is seen that the same H-bonds (with thymine of T18″, with the sugar unit of T19″, and with the phosphate of A21″) mediated *in trans* by K219 are lost by the tryptophan (W) substitution in the attacking subunit (Figure 5E). Interestingly, IntI1 with the mutation K219W mutant has previously been shown to bind strongly to ds *attI* but appears to be recombinationally deficient *in vivo* (*31*). We have observed that the K219W integrase binds very weakly to *attC*. This is a strong indication that K219 appears to be involved in hairpin binding and recognition, since only the *attC* site is proposed to form hairpins.

A majority of the mutated amino acids that were nonresponsive in EMSA are either situated at the protein surface or involved in hydrophobic packing. An exception is the tyrosine at position 220, which is in contact with bound DNA. It is situated in the C-terminus of α-helix I_2 and was mutated to either serine or phenylalanine. Serine was chosen because it is an aliphate with a hydroxyl group and phenylalanine is just a plain aromate, thus splitting the two features of the tyrosine into two separate clones. These Y mutations were nonresponsive in the EMSA, and Y220S integrase was also tested under denaturing conditions and found to be catalytically active (Figure 3A,B). These results were somewhat surprising since in our model Y220 in the attacking subunit makes a hydrogen bond to the phosphate between G20″ and A21″ and thereby stabilizes the G20″ bulge (Figure 5E). This H-bond is lost when Y220 is mutated to a fenylalanine and is apparently not essential for DNA binding. A serine in this position cannot make a hydrogen bond to DNA, but neither binding nor catalytic activity was affected. Apparently, a bulky aromatic ring in this position is not essential for DNA binding.

Can Different Requirements of IntI1 and VchIntIA for attC Recognition Explain the Broader Recognition Profile of IntI1? Those *attC* sites that have been found in integrons borne on plasmids are rather diverse (*9, 44, 45*). In contrast, *attC* sites on the chromosomes of many genera of gammaproteobacteria are conserved intergenic repeat families, homogeneous, and specific for each species (*6, 46*). However, even though all *attC* sites in the pool can vary considerably in length and sequence, they share the potential to form a stem−loop structure. It has been demonstrated that the range of *attC* sites efficiently recombined by VchIntIA is narrower than that of IntI1 (*36*). We have searched for similarities and dissimilarities between the two integrases that could further explain this and the unusual site recognition of integron integrases. Our observations from homology modeling of IntI1 against the VchIntIA−VCR_{bs} synaptic complex are summarized in Figure 6 where base-specific hydrogen bonding between VchIntIA−VCR_{bs} and IntI1−VCR_{bs} structures is compared. The MD simulations support the hydrogen bonds found in the IntI1 model.

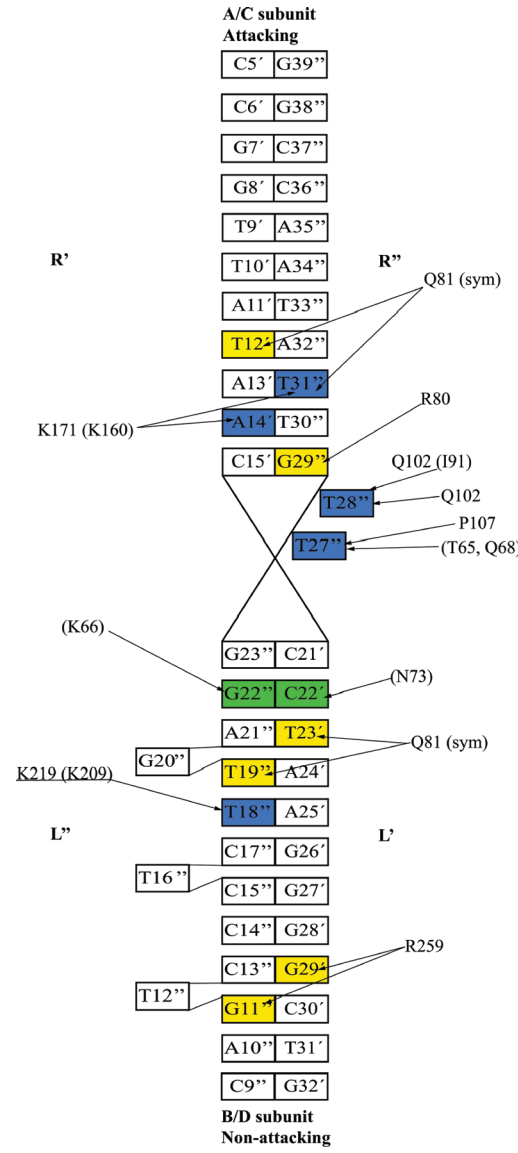

FIGURE 6: Base-specific hydrogen bonding presented earlier in the VchIntIA−VCR_{bs} complex (*32*) and in the homology model of the IntI1−VCR_{bs} complex from this work. Bases are shown as boxes, and the cross in the middle includes the bases in the spacer in VCR_{bs} in which only T28″ and T27″ are highlighted. The amino acids that form specific base contacts are indicated both in IntI1 and in VchIntIA (within parentheses for the latter). Those amino acids that interact in *trans* are underlined. Green boxes denote contacts only found with VchIntIA, blue boxes contacts found with both VchIntIA and IntI1, and yellow boxes bases only contacted by IntI1. The position of the loop in the natural hairpin conformation is downward.

Residues K66 and N73 in the nonattacking subunits of VchIntIA form hydrogen bonds to G22″ and C22′, respectively, that are lacking in IntI1. The two bases form a complementary pair in VCR_{bs}. Furthermore, residues T65 and Q68 both form a hydrogen bond to T27″. The base contacts made by the attacking subunits of VchIntIA are conserved in IntI1. Residue K171 in IntI1 (K160 in VchIntIA) forms hydrogen bonds in *cis* to bases A14′ and T31″ in VCR_{bs}. The backbone amide nitrogen of Q102 in IntI1 (I91 in VchIntIA) forms a hydrogen bond to T28″, also in *cis*. Residue K219 in IntI1 (K209 in VchIntIA) makes a hydrogen bond to T18″ in *trans*.

Different amino acids in the protein−DNA interface of the IntI1−VCR_{bs} complex compared to the VchIntIA−VCR_{bs} complex allow differences in sequence specificity. Due to such interface differences, nine possible base contacts not present in the VchIntIA structure are found in the IntI1 homology model. All these base contacts are made in *cis*. In both the attacking and nonattacking subunits of IntI1, Q81 forms hydrogen bonds to two diagonally positioned thymines in VCR_{bs} (T31″ and T12′ in attacking and T23′ and T19″ in nonattacking). In the attacking subunits, three additional hydrogen bonds are made through R80 to G29″, the side chain of Q102 to T28″, and the backbone carbonyl of P107 to T27″. In the nonattacking subunits, R259 contacts the two diagonally positioned residues, G11″ and G29′, that are situated on either side of the T12″ bulge.

DISCUSSION

The recently presented three-dimensional structure of the DNA-complexed *V. cholerae* integron integrase (*32*) is likely to be exceptionally important for unraveling the mechanism of gene cassette mobilization. In fact, the integron example widens the conception of how site-specific recombinases and DNA-binding proteins in general may recognize and process the target polymer. Other proteins with action directed against single-stranded DNA might share elements with the integron example, providing a connection to the paradigm set by the classical tyrosine recombinases like Cre and λ Int. The current integron integrase recombination model (*34*) suggests a single-stranded pathway that stops at the Holliday junction intermediate left to be resolved into the excision or integration product by yet unknown cellular mechanisms. A previous model, for which there is some published support (*10*), is based on recombination on the double-stranded level to follow a Campbell-like recombination trajectory. Another fact calling for a revised model is the wider site recognition profile of integron integrases compared to the typical tyrosine recombinases. Not only do the sites vary by sequence to an exceptional degree, integron integrases are furthermore capable of recombining two architecturally different site classes, *attI* and *attC*. It has been reported previously that integron integrases recognize the double-stranded form of *attI* (*47, 48*), a site in *cis* with the integrase gene. In contrast, and as discussed with respect to *attC*, binding occurs to a folded hairpin structure of the bottom strand, a prerequisite met by the palindromic appearance of this site class (*33, 49*). Remarkably, the top strand of the same sequence is not bound with the concomitant effect that only one end of the very long site recombines. Recognition of chiral bulges in the hairpin structure of *attC* enables recognition of the sites in spite of their sequence variation. The presented crystal structure (*32*) underlines the importance of the bulges for site recognition and points out amino acids that are directly involved in the interactions with these. In the crystal structure of four protomers of VchIntIA bound to two antiparallel VCR_{bs} duplexes, two subunits are attacking which means that they have the reactive tyrosine close to the DNA backbone. The activity could be regulated by positioning of the tyrosine in the four subunits. The structure supports a one-step cleavage and rejoining integrase-mediated process that stops at the Holliday junction stage, since the formed synapse has only a 2-fold symmetry and an isomerization step required for the second step of cleavage and rejoining would probably be energetically too costly.

The suggested single-stranded recombination pathway differs from the well-known two-step recombination mechanism utilized by numerous tyrosine recombinases. One theory is that the proposed unusual recombination mechanism of integron integrases is partly mediated by a unique insertion [16−36 amino acids in length (*31, 50, 51*)] in this subfamily of tyrosine recombinases. Amino acids in the unique region have now been demonstrated to be involved in interactions with the bulges and important for synapse formation (*32*). Interestingly, the location of two conserved prolines flanking αI_2 in this region indicates a specific role in recognition since they appear to limit the length of the helix by acting as helix breakers.

In our homology model of IntI1, we have investigated the environment around the catalytic residues and the residues involved in extrahelical base contacts with the important G20″ and T12″ bulges. One subaim of the modeling was to search the two proteins for features that could explain the broader recognition profile for IntI1 compared to that for VchIntIA. The *V. cholerae* enzyme, like all superintegron integrases, seems to have a preference for its own homogeneous *attC* sites, termed VCRs (ref *36* and references cited therein). We looked at base-specific hydrogen bonds between VchIntIA−VCR_{bs} and IntI1−VCR_{bs} complexes and found similar contacts but also contacts that differ between the two proteins. These interface differences could potentially explain why the recognition profile of IntI1 is suggested to be broader than that of VchIntIA. Interestingly, we observed more extensive base contacts in our homology model of IntI1 bound to VCR_{bs} compared to the VchIntIA−VCR_{bs} crystal structure. The specific base contacts made by IntI1 and shown in Figure 6 are at conserved positions among *attC* sites (*36*), suggesting that IntI1 to a greater extent compared to VchIntI1 allows variations elsewhere in the site and thus has a broader recognition profile. It has been reported that VchIntIA mediated recombination with ~2000-fold higher frequency in *V. cholerae* than in *E. coli* whereas IntI1 could perform recombination equally well under the two conditions (*36*). It is possible that host proteins play a role in full recombination and/or have regulatory functions *in vivo* especially since the last step in the recombination is thought to be mediated by cellular factors. That role could be of different importance in the two cases. Being a traveling integrase that is borne on mobile elements, IntI1 under evolution might have been selected in the direction of less restrictive recognition of *attC* sites in many different hosts.

In gel shift analysis, four mutations signaled either lost or greatly weakened binding to the bottom strand of *attC* compared to the distinct band seen in the EMSA for the wild-type integrase. The mutated integrase IntI1L216R yielded weaker binding to the oligonucleotide. This could according to our model be due to hydrophobic contacts between L216 and the G20″ bulge. The clear response to alterations with respect of the lysine at position 219 could mean that the version of the motif specific for the integron integrases has unique features. When the results from this work were compared with Messier and Roy's study (*31*) on the unique insertion region, an interesting fact appeared: the K219W integrase strongly binds to *attI* and appears to be recombinationally deficient. Here we have shown that the same

mutation results in a very weak binding of IntI1 to *attC*. This is a strong indication that K219 is involved in hairpin binding and hence in recognition of the folded *attC* site. The highly conserved tyrosine residue at position 220 did not affect binding at all. This was unexpected given the strong conservation of this amino acid but was reinforced by identical results with the two different mutations, Y220F and Y220S. One of the two mutations, Y220S, was tested under denaturing EMSA conditions and revealed a catalytic proficiency by forming a covalent linkage between its reactive tyrosine 312 and the DNA. It is possible that mutations in Y220 make the integrase inactive in recombination even though hairpin binding is not impaired.

In general through this and other recent studies, the picture of a functionally and structurally defined subclass of tyrosine recombinases is becoming clearer. Furthermore, recently published genomic data from wider ranges of microbes promise a clearer definition of the integron aspect of bacterial evolution (*52*), and data reporting a relation between *attC* and other recombination systems need to be addressed further (*53*). Most important in the immediate context of this study seems to be unraveling of the differential behavior of IntI proteins relative to its distinct classes of sites (*54*).

REFERENCES

1. Stokes, H. W., and Hall, R. M. (1989) A novel family of potentially mobile DNA elements encoding site-specific gene-integration functions: Integrons. *Mol. Microbiol. 3*, 1669–1683.
2. Hall, R. M., and Collis, C. M. (1995) Mobile gene cassettes and integrons: Capture and spread of genes by site-specific recombination. *Mol. Microbiol. 15*, 593–600.
3. Leverstein-van Hall, M. A., Blok, H. E. M., Donders, A. R. T., Paauw, A., Fluit, A. C., and Verhoef, J. (2003) Multidrug resistance among Enterobacteriaceae is strongly associated with the presence of integrons and is independent of species or isolate origin. *J. Infect. Dis. 187*, 251–259.
4. Martinez, E., and de la Cruz, F. (1990) Genetic elements involved in Tn21 site-specific integration, a novel mechanism for the dissemination of antibiotic resistance genes. *EMBO J. 9*, 1275–1281.
5. Sundstrom, L., Radstrom, P., Swedberg, G., and Skold, O. (1988) Site-specific recombination promotes linkage between trimethoprim and sulfonamide resistance genes. Sequence characterization of dhfrV and sulI and a recombination active locus of Tn21. *Mol. Gen. Genet. 213*, 191–201.
6. Mazel, D., Dychinco, B., Webb, V. A., and Davies, J. (1998) A distinctive class of integron in the *Vibrio cholerae* genome. *Science 280*, 605–608.
7. Burrus, V., and Waldor, M. K. (2004) Shaping bacterial genomes with integrative and conjugative elements. *Res. Microbiol. 155*, 376–386.
8. Sundstrom, L., Roy, P. H., and Skold, O. (1991) Site-specific insertion of three structural gene cassettes in transposon Tn7. *J. Bacteriol. 173*, 3025–3028.
9. Recchia, G. D., and Hall, R. M. (1995) Gene cassettes: A new class of mobile element. *Microbiology (Reading, U.K.) 141*, 3015–3027.
10. Collis, C. M., Grammaticopoulos, G., Briton, J., Stokes, H. W., and Hall, R. M. (1993) Site-specific insertion of gene cassettes into integrons. *Mol. Microbiol. 9*, 41–52.
11. Collis, C. M., and Hall, R. M. (1992) Gene cassettes from the insert region of integrons are excised as covalently closed circles. *Mol. Microbiol. 6*, 2875–2885.
12. Collis, C. M., and Hall, R. M. (1992) Site-specific deletion and rearrangement of integron insert genes catalyzed by the integron DNA integrase. *J. Bacteriol. 174*, 1574–1585.
13. Grindley, N. D. F., Whiteson, K. L., and Rice, P. A. (2006) Mechanisms of Site-Specific Recombination. *Annu. Rev. Biochem. 75*, 567–605.
14. Nash, H. A. (1996) Site-specific recombination: Integration, excision, resolution, and inversion of defined DNA segments. In *Escherichia coli and Salmonella: Cellular and Molecular Biology* (Neidhardt, F. C., Curtiss, R., III, Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., Eds.) 2nd ed., pp 2363−2376, ASM Press, Washington, DC.
15. Mondragon, A. (1997) Solving the cis/trans paradox in the Int family of recombinases. *Nat. Struct. Biol. 4*, 427–429.
16. Chaconas, G., Stewart, P., Tilly, K., Bono, J., and Rosa, P. (2001) Telomere resolution in the Lyme disease spirochete. *EMBO J. 20*, 3229–3237.
17. Argos, P., Landy, A., Abremski, K., Egan, J. B., Haggard-Ljungquist, E., Hoess, R. H., Kahn, M. L., Kalionis, B., Narayana, S. V., and Pierson, L. S. (1986) The integrase family of site-specific recombinases: Regional similarities and global diversity. *EMBO J. 5*, 433–440.
18. Nunes-Duby, S. E., Kwon, H. J., Tirumalai, R. S., Ellenberger, T., and Landy, A. (1998) Similarities and differences among 105 members of the Int family of site-specific recombinases. *Nucleic Acids Res. 26*, 391–406.
19. Esposito, D., and Scocca, J. J. (1997) The integrase family of tyrosine recombinases: Evolution of a conserved active site domain. *Nucleic Acids Res. 25*, 3605–3614.
20. Abremski, K. E., and Hoess, R. H. (1992) Evidence for a second conserved arginine residue in the integrase family of recombination proteins. *Protein Eng. 5*, 87–91.
21. Grainge, I., and Jayaram, M. (1999) The integrase family of recombinase: Organization and function of the active site. *Mol. Microbiol. 33*, 449–456.
22. Yang, W., and Mizuuchi, K. (1997) Site-specific recombination in plane view. *Structure 5*, 1401–1406.
23. Guo, F., Gopaul, D. N., and van Duyne, G. D. (1997) Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature 389*, 40–46.
24. Chen, Y., Narendra, U., Iype, L. E., Cox, M. M., and Rice, P. A. (2000) Crystal structure of a Flp recombinase-Holliday junction complex: Assembly of an active oligomer by helix swapping. *Mol. Cell 6*, 885–897.
25. Biswas, T., Aihara, H., Radman-Livaja, M., Filman, D., Landy, A., and Ellenberger, T. (2005) A structural basis for allosteric control of DNA recombination by lambda integrase. *Nature 435*, 1059–1066.
26. Redinbo, M. R., Stewart, L., Kuhn, P., Champoux, J. J., and Hol, W. G. (1998) Crystal structures of human topoisomerase I in covalent and noncovalent complexes with DNA. *Science 279*, 1504–1513.
27. Kwon, H. J., Tirumalai, R., Landy, A., and Ellenberger, T. (1997) Flexibility in DNA recombination: Structure of the lambda integrase catalytic core. *Science 276*, 126–131.
28. Hickman, A. B., Waninger, S., Scocca, J. J., and Dyda, F. (1997) Molecular organization in site-specific recombination: The catalytic domain of bacteriophage HP1 integrase at 2.7 Å resolution. *Cell 89*, 227–237.
29. Subramanya, H. S., Arciszewska, L. K., Baker, R. A., Bird, L. E., Sherratt, D. J., and Wigley, D. B. (1997) Crystal structure of the site-specific recombinase, XerD. *EMBO J. 16*, 5178–5187.
30. Sharma, A., Hanai, R., and Mondragon, A. (1994) Crystal structure of the amino-terminal fragment of vaccinia virus DNA topoisomerase I at 1.6 Å resolution. *Structure 2*, 767–777.
31. Messier, N., and Roy, P. H. (2001) Integron integrases possess a unique additional domain necessary for activity. *J. Bacteriol. 183*, 6699–6706.
32. MacDonald, D., Demarre, G., Bouvier, M., Mazel, D., and Gopaul, D. N. (2006) Structural basis for broad DNA-specificity in integron recombination. *Nature 440*, 1157–1162.
33. Johansson, C., Kamali-Moghaddam, M., and Sundstrom, L. (2004) Integron integrase binds to bulged hairpin DNA. *Nucleic Acids Res. 32*, 4033–4043.
34. Bouvier, M., Demarre, G., and Mazel, D. (2005) Integron cassette insertion: A recombination process involving a folded single strand substrate. *EMBO J. 24*, 4356–4367.
35. Rowe-Magnus, D. A., Guerout, A. M., and Mazel, D. (2002) Bacterial resistance evolution by recruitment of super-integron gene cassettes. *Mol. Microbiol. 43*, 1657–1669.
36. Biskri, L., Bouvier, M., Guerout, A. M., Boisnard, S., and Mazel, D. (2005) Comparative study of class 1 integron and *Vibrio cholerae* superintegron integrase activities. *J. Bacteriol. 187*, 1740–1750.

37. Notredame, C., Higgins, D. G., and Heringa, J. (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol. 302*, 205–217.
38. Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. (2003) SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res. 31*, 3381–3385.
39. Marelius, J., Kolmodin, K., Feierberg, I., and Aqvist, J. (1998) Q: A molecular dynamics program for free energy calculations and empirical valence bond simulations in biomolecular systems. *J. Mol. Graphics Modell. 16*, 213–225, 261.
40. MacKerell, A. D., Jr., Bashford, D., Bellott, M., Dunbrack, R. L., Jr., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., III, Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J. M., Watanabe, M., Wiórkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B 102*, 3586–3616.
41. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential function for simulating liquid water. *J. Chem. Phys. 79*, 926–935.
42. King, G., and Warshel, A. (1989) A surface constrained all-atom solvent model for effective simulations of polar solutions. *J. Chem. Phys. 91*, 3647–3661.
43. Lee, F. S., and Warshel, A. (1992) A local reaction field method for fast evaluation of long-range electrostatic interactions in molecular simulations. *J. Chem. Phys. 97*, 3100–3107.
44. Francia, M. V., Avila, P., de la Cruz, F., and Garcia Lobo, J. M. (1997) A hot spot in plasmid F for site-specific recombination mediated by Tn21 integron integrase. *J. Bacteriol. 179*, 4419–4425.
45. Stokes, H. W., O'Gorman, D. B., Recchia, G. D., Parsekhian, M., and Hall, R. M. (1997) Structure and function of 59-base element recombination sites associated with mobile gene cassettes. *Mol. Microbiol. 26*, 731–745.
46. Barker, A., Clark, C. A., and Manning, P. A. (1994) Identification of VCR, a repeated sequence associated with a locus encoding a hemagglutinin in *Vibrio cholerae* O1. *J. Bacteriol. 176*, 5450–5458.
47. Collis, C. M., Kim, M. J., Stokes, H. W., and Hall, R. M. (1998) Binding of the purified integron DNA integrase Intl1 to integron- and cassette-associated recombination sites. *Mol. Microbiol. 29*, 477–490.
48. Gravel, A., Fournier, B., and Roy, P. H. (1998) DNA complexes obtained with the integron integrase IntI1 at the attI1 site. *Nucleic Acids Res. 26*, 4347–4355.
49. Francia, M. V., Zabala, J. C., de la Cruz, F., and Garcia Lobo, J. M. (1999) The IntI1 integron integrase preferentially binds single-stranded DNA of the attC site. *J. Bacteriol. 181*, 6844–6849.
50. Nield, B. S., Holmes, A. J., Gillings, M. R., Recchia, G. D., Mabbutt, B. C., Nevalainen, K. M., and Stokes, H. W. (2001) Recovery of new integron classes from environmental DNA. *FEMS Microbiol. Lett. 195*, 59–65.
51. Recchia, G. D., and Sherratt, D. J. (2002) Gene acquisition in bacteria by integron-mediated site-specific recombination. In *Mobile DNA II* (Craig, N. L., Craigie, R., Gellert, M., and Lambowitz, A. M., Eds.) 2nd ed., pp 162−176, ASM Press, Washington, DC.
52. Gillings, M., Boucher, Y., Labbate, M., Holmes, A., Krishnan, S., Holley, M., and Stokes, H. W. (2008) The Evolution of Class 1 Integrons and the Rise of Antibiotic Resistance. *J. Bacteriol. 190*, 5095–5100.
53. Quiroga, C., Roy, P. H., and Centrón, D. (2008) The S.ma.I2 class C group II intron inserts at integron attC sites. *Microbiology (Reading, U.K.) 154*, 1341–1353.
54. Demarre, G., Frumerie, C., Gopaul, D. N., and Mazel, D. (2007) Identification of key structural determinants of the IntI1 integron integrase that influence attC × attI1 recombination efficiency. *Nucleic Acids Res. 35*, 6475–6489.
55. Gravel, A., Messier, N., and Roy, P. H. (1998) Point mutations in the integron integrase IntI1 that affect recombination and/or substrate recognition. *J. Bacteriol. 180*, 5437–5442.

BI8020235