JOURNAL OF BACTERIOLOGY, Apr. 2009, p. 2330–2339 Vol. 191, No. 7
0021-9193/09/$08.00+0 doi:10.1128/JB.01280-08

Mutational Analysis and Homology-Based Modeling of the IntDOT Core-Binding Domain▽

Karolina Malanowska,[1] Joel Cioni,[1] Brian M. Swalla,[2] Abigail Salyers,[1] and Jeffrey F. Gardner[1]*

Department of Microbiology, University of Illinois, Urbana, Illinois 61801,[1] and Athena Biotechnologies, Inc., Newark, Delaware 19711[2]

Received 11 September 2008/Accepted 14 January 2009

Tyrosine recombinases mediate a wide range of important genetic rearrangement reactions. Models for tyrosine recombinases have been based largely on work done on the integrase of phage lambda and recombinases like Cre, Flp, and XerC/D. All of these recombinases share a common amino acid signature that is important for catalysis. Several conjugative transposons (CTns) encode recombinases that are also members of the tyrosine recombinase family, but the reaction that they catalyze differs in that recombination does not require homology in the attachment sites. In this study, we examine the role of the core-binding (CB) domain of the CTnDOT integrase (IntDOT) that is located adjacent to the catalytic domain of the protein. Since there is no crystal structure for any of the CTn integrases, we began with a predicted three-dimensional structure produced by homology-based modeling. Amino acid substitutions were made at positions predicted by the model to be close to the DNA. Mutant proteins were tested for the ability to mediate integration in vivo and for in vitro DNA-binding, cleavage, and ligation activities. We identified for the first time nonconserved amino acid residues in the CB domain that are important for catalytic activity. Mutant proteins with substitutions at three positions in the CB domain are defective for DNA cleavage but still proficient in ligation. The positions of the residues in the complex suggest that the mutant residues affect the positioning of the cleaved phosphodiester bond in the active site without disruption of the ligation step.

Conjugative transposons (CTns), also known as integrative and conjugative elements, are DNA elements that can excise from the bacterial chromosome as a double-stranded circle and transfer by conjugation to another bacterial host, where they integrate into the recipient chromosome. CTnDOT is an example of a CTn that resides in the *Bacteroides* chromosome. The integrase of CTnDOT (IntDOT), which is essential for both excision and integration, was classified as a member of the tyrosine recombinase family (17) based on identification of five of the six active site amino acids that are conserved in tyrosine recombinases, RK(H/K)R(H/W)Y, which perform the chemistry involved in the cleavage and ligation during CTnDOT integration. One of the essential amino acids was the predicted catalytic tyrosine. Previously, we demonstrated that changing three of the six conserved amino acids eliminated integration of IntDOT in vivo and abolished or reduced cleavage and ligation activity in vitro (17). In other tyrosine recombinases, changes in all six of these conserved amino acids eliminated or drastically reduced recombination (for a review, see reference 29). Our mutational analysis of the catalytic region of IntDOT suggested that the architecture of the active site is similar to that of other tyrosine recombinases, such as the integrase of phage λ (λ Int), Cre, Flp, and XerC/D. However, IntDOT-catalyzed recombination is significantly different because it is not site specific and tolerates heterology in the crossover region between the CTnDOT *attDOT* site and the chromosomal *attB* site (5, 7).

Tyrosine recombinases have two or three domains, depending upon whether the system includes regulated integration and excision reactions. The C-terminal catalytic domain contains all residues required for catalysis. The core-binding (CB) domain interacts primarily with the major groove on the attachment site and facilitates binding to the core DNA sequence (2, 29). Some recombinases, like λ Int and IntDOT, have an additional amino-terminal domain that is poorly conserved throughout the family (2, 17, 29). Biochemical experiments performed in different laboratories have suggested that this domain recognizes sites, called arm-type sites, that flank the crossover region and gives directionality to the recombination reaction (4, 25). Accessory proteins, like integration host factor (IHF), also bind to regions flanking the core-type sites and bend the DNA to promote the formation of higher-order nucleoprotein complexes called intasomes that are required for synapsis and recombination (3, 20, 27).

A useful method for assessing the similarity between IntDOT and previously studied tyrosine recombinases is structural comparison. Recombinases such as λ Int and Cre have structural similarities even though their primary amino acid sequences are very different (26, 29). Since IntDOT has not yet been crystallized, we obtained structural information through homology-based modeling to predict the three-dimensional structure of the CB domain. With the exception of the work described above characterizing active site residues, a mutational study has not been performed for IntDOT. Accordingly, we used the predicted three-dimensional structure to identify and study the effects of site-directed mutations at amino acid positions predicted to interact with attachment site DNA. We show here that a useful predicted structure for the IntDOT CB domain (Fig. 1) can be constructed from the known structures

* Corresponding author. Mailing address: Department of Microbiology, B103 Chemical and Life Sciences Laboratory, 601 South Goodwin Avenue, University of Illinois, Urbana, IL 61801. Phone: (217) 333-7287. Fax: (217) 244-6697. E-mail: jeffgard@uiuc.edu.

▽ Published ahead of print on 23 January 2009.

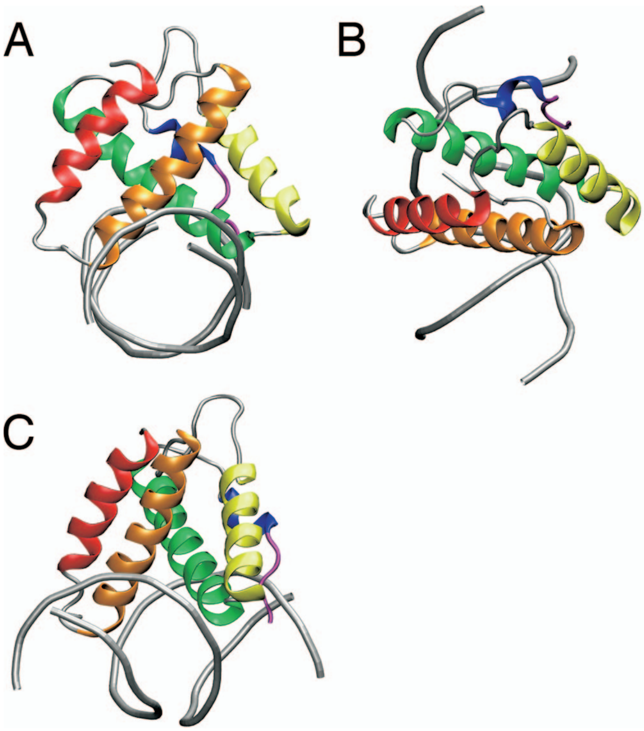

FIG. 1. Cartoon representation of the three-dimensional homology model of the IntDOT CB domain bound to DNA. The model is based on the cocrystal structures of Cre and λ Int bound to a core-type DNA site. The α-helices of the CB domain are indicated as follows: red, helix 1; orange, helix 2; yellow, helix 3; green, helix 4; and blue, helix 5. The backbone of the bound DNA helix is gray. The orthogonally crossed helices, helix 2 (orange) and helix 4 (green), penetrate the major groove of the DNA. Panels A, B, and C show three different representations of the model, in which the structure is viewed in different orientations.

of other well-studied tyrosine recombinases (9, 10, 26) due to their conserved structural features. We used our model to guide biochemical experiments that demonstrated a role for the IntDOT CB domain in the catalytic reaction.

MATERIALS AND METHODS

Construction of a homology-based model for IntDOT. A three-dimensional model of the IntDOT CB domain was created using methods described previously for modeling of the λ Int CB domain (26). This approach produced a λ Int CB domain model that was subsequently confirmed by a crystal structure for this domain (1). To model the IntDOT CB domain, the mGenTHREADER fold recognition algorithm (13, 19) was employed to identify structures in the Protein Data Bank showing structural homology with the IntDOT CB sequence. Significant homology was identified for the CB domain, which aligned with Cre (4crx), XerD (1a0p), and λ Int (1p7d) (Fig. 2) (1, 9, 10, 24). The sequence-structure alignments produced by mGenTHREADER were confirmed through comparison with alignments produced by other methods, including both threading-based (3D-PSSM, ORFeus-2, BasicD, INGBU, and FUGUE2) and sequence-based (FFAS03 and Psi-BLAST) alignment algorithms. Construction of a multiple-sequence alignment showed that each individual alignment generally agreed with the results obtained initially with mGenTHREADER, and the mGenTHREADER alignment was thus selected to guide the model-building process. Each recombinase structure was evaluated for its suitability as a template by manually threading the IntDOT sequence onto each template using Deep View Swiss-PdbViewer software to match the sequence-structure alignment produced by mGenTHREADER. Each template was evaluated by inspecting the threading energy of the IntDOT sequence at each amino acid position as calculated by Swiss-PdbViewer. Overall, the Cre 4crx structure provided the best fit and was judged to be the most suitable template; however, certain high-energy regions (within IntDOT residues V112 to K129 and C181 to E219) were observed, in which the threading could be improved. These specific problem areas corresponded to surface residues and "loop" regions between secondary structure elements for which weak structural conservation is anticipated. The threading energy of these regions was improved by employing the λ Int (1ae9) crystal structure as the template. A hybrid template was then constructed, in which the suboptimal regions of the 4crx template were replaced with the corresponding atomic coordinates from the 1ae9 structure. The hybrid template was constructed essentially as described previously (26). To build the model, the IntDOT sequence was manually threaded onto the hybrid template as described above, and the model was constructed using the Swiss-Model server (22). The model-building process employed by Swiss-Model incorporates a molecular dynamics simulation to minimize the energy of backbone and side chain conformations and a loop-building process to predict the best position and orientation for residues in flexible protein regions that are not well conserved. These calculations were performed using the default parameters. After construction, the model was examined using PROCHECK to inspect and validate protein structural properties, such as residue geometry and conformation.

Bacterial strains, plasmids, and growth conditions. The strains and plasmids used in this study are listed in Table 1. *Escherichia coli* strains were grown in Luria-Bertani (LB) medium. *Bacteroides* strains were grown in chopped meat medium (Remel) and in Trypticase yeast extract glucose (TYG) medium. All


```
A.  IntDOT predicted secondary structure:
                110       120       130       140       150       160       170       180       190       200       210
IntDOT:   LEHRYHTLMQVFRRHNEDYEKQVEAGMKAKGTLLKYRTVYKHLQEFLTIRYHVKDIALKELTPAFISDFEMFLRTDKHCCTNTVWLYVCPLRTMVFIAINNEWLTRDPFREYEI
Pred:     CCCCCHHHHHHHHHHHHHHHHHHCCCHHHHHHHHHHHHHHHHHHHHHHHCCCCCCCHHHCCHHHHHHHHHHHHHHHHCCCHHHHHHHHHHHHHHHHHHHHCCCCCCCCCCCCCC
Conf:     997632799999999999999978670119989999999999999999862587788489769999999999988108788799999999999999999879965587050201

B.  Alignment between IntDOT and Cre: 9.9% identical (9/91)
                          10          20        30        40            50        60         70         80         90
Cre SS:   --------CHHHHHHHHHHHH--CCCCCCHHHHHHHHHHHHHHHHHHHHH----CCCCCCCCHHHHHHHHHHHH-HCCCCHHHHHHHHHHHHHHHHCCCCCCCCCC--------
Cre:      --------SDEVRKNLMDMFR--DRQAFSEHTWKMLLSVCRSWAAWCKLN----NRKWFPAEPEDVRDYLLYLQ-ARGLAVKTIQQHLGQLNMLHRRSGLPRPSDS--------
                      |    |             |      |                       |    |    |         |       |
IntDOT:   LEHRYHTLMQVFRRHNEDYEKQVEAGMKAKGTLLKYRTVYKHLQEFLTIRYHVKDIALKELTPAFISDFEMFLRTDKHCCTNTVWLYVCPLRTMVFIAINNEWLTRDPFREYEI

C.  Alignment between IntDOT and Lambda Int: 18.3% identical (17/93)
                        80        90        100      110              120       130         140       150       160
Int SS:   ---------CCHHHHHHHHHHHHHHHCCCHHHHHHHHHHHHHHHHHC------CCCCCCCCCHHHHHHHHHHHHHCCC--HHHHHHHHHHHHHHHHHHHHCCCCCCCCCC----
Int:      ---------VTLHSWLDRYEKILASRGIKQKTLINYMSKIKAIRRGL------PDAPLEDITTKEIAAMLNGYIDEGK--AASAKLIRSTLSDAFREAIAEGHITTNHVA----
                            |||          ||  |    |     |       |  |   |   |                   |    |      ||     |
IntDOT:   LEHRYHTLMQVFRRHNEDYEKQVEAGMKAKGTLLKYRTVYKHLQEFLTIRYHVKDIALKELTPAFISDFEMFLRTDKHCCTNTVWLYVCPLRTMVFIAINNEWLTRDPFREYEI
```

FIG. 2. (A) Secondary structure of IntDOT predicted by PSIPRED (14). C, coil; H, alpha-helix. The confidence level of the prediction at each position is indicated using a scale from 0 (low) to 9 (high). (B and C) Alignment of IntDOT with Cre and λ Int (Int) used to produce the model described in this study. The secondary structure observed in each crystal structure is indicated on the SS line. Positions with identical residues are indicated by vertical lines.

TABLE 1. Bacterial strains and plasmids

Strain or plasmid(s)	Relevant phenotype[a]	Description	Source and/or reference
E. coli strains			
DH5αMCR	$RecA^-$		GibcoBRL
BL21(DE3)/pLysE	(DE3)		Invitrogen (23)
S17-1	$RecA^-$ Tmp^r Str^r	Plasmid RP4 inserted into the S17-1 chromosome to supply transfer functions	5
Bacteroides thetaiotaomicron BT 4001	Rif^r Gen^r	Spontaneous rifampin-resistant mutant of *Bacteroides thetaiotaomicron* 5482A	5
Plasmids			
pDJE2.3	Amp^r Erm^r		
pDJE2.3 derivatives	Amp^r Erm^r Kan^r	Isogenic plasmids that contain substitution mutations in the *intDOT* gene	This study
pET30	Kan^r	Vector containing T7 promoter for overexpression of proteins	23
pET30a derivatives	Kan^r		This study

[a] Abbreviations: Amp^r, ampicillin resistance; Erm^r, erythromycin resistance; Kan^r, kanamycin resistance; Tmp^r, trimethoprim resistance; Str^r, streptomycin resistance; Gen^r, gentamicin resistance.

antibiotics were supplied by Sigma. Restriction and DNA modification enzymes were supplied by New England Biolabs. *Pfu* and *Taq* polymerases were supplied by Stratagene and Invitrogen, respectively.

Site-directed mutagenesis of *intDOT*. The *intDOT* mutants used for the in vivo integration assays were made using pDJE2.3 as a template. The pDJE2.3 plasmid contains *attDOT*, the native *intDOT* gene, an ampicillin resistance gene (*bla*) that is expressed in *E. coli*, and an erythromycin resistance gene (*ermG*) that is expressed in *Bacteroides*. Using a Stratagene QuikChange mutagenesis kit, 15 *intDOT* mutants were generated: K129A, K131A, T133A, L135A, K136A, Y137A, R138A, T139A, K142A, H143A, H179A, C180A, N183A, W186A, and T194A. Primers used to generate the mutations are shown in Table 2. After mutagenesis was performed, the mutagenized plasmid was transformed into *E. coli* DH5αMCR competent cells. The cells were plated on LB containing ampicillin (100 μg/ml), and the plasmids were isolated from several independent colonies. The *intDOT* gene for several candidates of each mutant was sequenced to confirm the presence of the desired mutation and to ensure that there were no additional mutations.

The wild-type and mutant genes were then moved into the pET28 or pET30 vector as described previously (7) using primers I and II for PCRs. The primers used to PCR amplify the genes generate NdeI and HindIII ends. The PCR products and pET30a vector were treated with NdeI and HindIII, ligated together, and electroporated into competent DH5αMCR cells. Constructs with

TABLE 2. Oligonucleotides used for site-directed mutagenesis

Oligonucleotide	Sequence	Substitution in IntDOT
FrvK129A	CAGGTGGAGGCTGGCATGGCGGCCAAAGGCACGCTGTTG	K129A
BcvK129A	CAACAGCGTGCCTTTGGCCGCCATGCCAGCCTCCACCTG	K129A
FrvK131A	GGTACTTCAACAGCGTGCCCGCGGCCTTTCATGCCAGCCTCCACCT	K131A
BcvK131A	CAGGTGGAGGCTGGCATGAAAGCCGCGGGCACGCTGTTGAAGTACC	K131A
FrvT133A	GGCATGAAAGCCAAAGGCGCGCTGTTGAAGTACCGCACC	T133A
BcvT133A	GGTGCGGTACTTCAACAGCGCGCCTTTGGCTTTCATGCC	T133A
FrvL135A	ATGAAAGCCAAAGGCACGCTGGCGAAGTACCGCAACGTTTAC	L135A
BcvL135A	GTAAACGGTGCGGTACTTCGCCAGCGTGCCTTTGGCTTTCAT	L135A
FrvK136A	GGCAAAGGCACGCTGTTGGCGTACCGCACCGTTTACAAGC	K136A
BcvK136A	GCTTGTAAACGGTGCGGTACGCCAACAGCGTGCCTTTGCC	K136A
FrvY137A	GGCAGCCTGTTGAAGGCCCGCACCGTTTACAAG	Y137A
BcvY137A	CTTGTAAACGGTGCGGGCCTTCAACGGCGTGCC	Y137A
FrvR138A	GGCACGCTGTTGAAGTACGCCACCGTTTACAAG	R138A
BcvR138A	CTTGTAAACGGTGGCGTAGTTCAACGGCGTGCC	R138A
FrvT139A	CTGTTGAAGTACCGCGCCGTTTACAAGCACCTG	T139A
BcvT139A	CAGGTGCTTGTAAACGGCGCGGTACTTCAACAG	T139A
FrvK142A	GTACCGCACCGTTTACGCCCACCTGCAAGAGTT	K142A
BcvK142A	GAACTCTTGCAGGTGGGCGTAAACGGTGCGGTA	K142A
FrvH143A	CGCACCGTTTACAAGGCCCTGCAAGAGTTCCTC	H143A
BcvH143A	GAGGAACTCTTGCAGGGCCTTGTAAACGGTGCG	H143A
FrvH179A	CTGCGCACGGACAAGGCCTGCTGCACCAATACC	H179A
BcvH179A	GGTATTGGTGCAGCAGGCCTTGTCCGTGCGCAG	H179A
FrvC180A	CGCACGGACAAGCACGCCTGCACCAATACCGTG	C180A
BcvC180A	CACGGTATTGGTGCAGGCGTGCTTGTCCGTGCG	C180A
FrvN183A	CAAGCACTGCTGCACCGCCACCGTGTGGCTGTA	N183A
BcvY183A	TACAGCCACACGGTGGCGGTGCAGCAGTGCTTG	Y183A
FrvW186A	CTGCACCAATACCGTGGCCCTGTACGTCTGCCCG	W186A
BcvW186A	CGGGCAGACGTACAGGGCCACGGTATTGGTGCAG	W186A
FrvT194A	GTCTGCCCGCTACGGGCCCTGGTGTTCATCGCC	T194A
BcvT194A	GGCGATGAACACCATGGCCCGTAGCGGGCAGAC	T194A

insertions were then subjected to DNA sequence analysis to confirm that no PCR-induced mutations were present. The *intDOT* genes of these constructs are expressed by the T7 promoter.

In vivo integration assay. The S17-1 *E. coli* donor strain that was used for the mating experiments carried either the pDJE2.3 or pDJE2.3 plasmid containing mutant *intDOT* genes (5). The positive control was pDJE2.3 (wild-type *intDOT* gene), and the negative control was pDJE1.1, which contains a truncated version of *intDOT* (5). The integration frequencies were calculated by determining the ratio of recombinants to recipients. DH5αMCR donor strains were grown in LB containing ampicillin (100 μg/ml), and recipient strain BT4001 was grown in TYG medium (5). The cultures were allowed to grow to an optical density at 650 nm of 0.2. The donor and recipient cultures were mixed, collected by centrifugation (12,000 × *g* for 10 min 10°C), and then resuspended in 100 μl of TYG medium and placed on a nitrocellulose filter disk on a TYG agar plate. Plates were incubated aerobically overnight at 37°C. The cells on the filters were resuspended with TYG medium containing gentamicin (Gen) (200 μg/ml), diluted in TYG medium containing Gen, and plated on TYG medium containing Gen and on TYG medium containing Gen and erythromycin (10 μg/ml) to select for recipients and recombinants, respectively. The plates were incubated at 37°C anaerobically in GasPak jars for 2 days.

Protein purification. *E. coli* BL21(DE3)/pLysE containing the pET30a plasmid with the wild-type or mutant *intDOT* genes was grown overnight in LB medium supplemented with kanamycin (50 μg/ml) and chloramphenicol (20 μg/ml) at 30°C. The next day the culture was diluted 1:100 into LB medium (500 ml) supplemented with kanamycin and chloramphenicol and grown at 30°C until the optical density at 600 nm reached 0.5. Isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM, and the culture was grown at room temperature for an additional 4 to 5 h. Cells were harvested by centrifugation at 5,000 × *g* and stored at −80°C until they were used. The cell pellet was resuspended into 20 ml of lysis buffer (50 mM $NaHPO_4$ [pH 7], 50 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol [DTT], 5% glycerol, lysozyme [1 mg/ml], phenylmethylsulfonyl fluoride [50 μg/ml], one tablet of complete EDTA-free protease inhibitor [Roche]). The cell suspension was sonicated and centrifuged 27, 000 × *g* for 30 min at 4°C. The crude extract was centrifuged, and the supernatant was loaded onto a series of four 5-ml heparin Sepharose columns (GE Healthcare). The columns were washed with 5 column volumes of low-salt buffer (50 mM $NaHPO_4$ [pH 7], 50 mM NaCl, 1 mM EDTA, 1 mM DTT, 5% glycerol). The protein was eluted with 10 column volumes using a linear gradient from 50 mM to 1 M NaCl. IntDOT eluted at approximately 0.4 M NaCl. The cleavage assay was used to identify active fractions. Active fractions were subjected to sodium dodecyl sulfate (SDS) gel electrophoresis, in which IntDOT migrated as a 48-kDa band. The most active fractions were pooled and dialyzed into storage buffer (low-salt buffer containing 40% glycerol) overnight at 4°C.

DNA binding assay. A 320-bp *attDOT* DNA fragment was synthesized by PCR amplification using primers DRJ/MM160F (5′-TCGGGCATGGTTACGAAC-3′) and DLJ/AG161R (5′-CTCGAATTAAAATAGCTACTTTTG-3′). The PCR product was 5′ end labeled with [γ-^{32}P]ATP (6,000 Ci/mmol; Perkin Elmer) using T4 polynucleotide kinase (37°C, 1 h). The end-labeled DNA (12 nM) and IHF (20 nM) were incubated at room temperature for 10 min in 10 μl of reaction buffer (50 mM Tris-HCl [pH 8], 50 mM NaCl, 1 mM EDTA, 0.25 mg/ml bovine serum albumin [BSA], 75 ng/ml herring sperm DNA, 10% glycerol) to form a complex of IHF and *attDOT* DNA. Dilutions of the partially purified wild-type IntDOT or variants were incubated with the labeled DNA and IHF, and the samples were electrophoresed on a 5% nondenaturing polyacrylamide gel in 0.5× TBE buffer (0.0445 M Tris base, 0.0445 M borate, 0.001 M EDTA) for 3 h at 200 V. The gels were exposed to phosphorimager screens, and the levels of radiation in unbound and shifted DNA were estimated using a Fuji FLA-3000 phosphorimager and Fujifilm Image Gauge software (Macintosh v.3.4). A 1/32 dilution of the wild-type protein or a 1/16 dilution of the mutant protein preparations shifted one-half of the IHF-*attDOT* complexes to a more slowly migrating species.

Cleavage assay. Cleavage assays were performed as described by Malanowska et al. (17). The cleavage substrate consisted of the complete bottom strand of the *attDOT* core region (*attDOT-B*) annealed to a shorter top strand ending 3 bp downstream of the cleavage and strand exchange site (*attDOT-T4*). If IntDOT performs the cleavage step of the recombination reaction, the small 3-bp fragment diffuses away from the intermediate, preventing religation and leaving IntDOT covalently bound to the substrate. With a labeled substrate this protein-DNA complex can be seen as a band migrating more slowly than the substrate-only band on a polyacrylamide gel. After the top strand was labeled with ^{32}P, the top and bottom strands were annealed at a 1:10 ratio in annealing buffer (0.2 M NaCl in Tris-EDTA). The cleavage substrate (43 pmol) was then incubated with the amount of wild-type or mutant IntDOT that shifted one-half of the IHF-*attDOT* complex (see above) for 2 h at 37°C in the reaction buffer (20 mM Tris-HCl [pH 8],5 mM DTT, 0.05 mg/ml BSA, 1% glycerol). Control samples were also prepared by incubating the same reaction mixtures with proteinase K for an additional 2 h at 42°C to prove that the protein was covalently bound to the substrate. All samples were boiled for 5 min in Tris-glycerol-SDS sample buffer (63 mM Tris-HCl, 10% glycerol, 2% SDS, 0.0025% bromophenol blue; pH 6.8) and immediately loaded onto a 4 to 20% Tris-glycine gradient gel. Electrophoresis was performed at 100 V for 90 min. The assays were performed at least in triplicate. The gels were then exposed to phosphorimager screens and visualized using the Fuji FLA-3000 phosphorimager and Fujifilm Image Gauge software (Macintosh v.3.4).

Ligation assay. The substrate (17) used for the ligation assay was double-stranded *attDOT* oligonucleotide that contained a 3′ *para*-nitrophenol (pNP)-activated nick at the site of strand cleavage in the bottom strand. The 3′ pNP nick mimics the 3′ phosphotyrosine intermediate of recombination just after the cleavage step (31). The 5′-end-labeled bottom strand was mixed with the other bottom strand (JG02B) and top strand (JG02T) at a ratio of 1:5:10 and annealed in annealing buffer (0.2 M NaCl in Tris-EDTA) by heating the preparation to 95°C and cooling it at a rate of 1°C/min to 22°C. The ligation reactions took place in ligation buffer (50 mM Tris-HCl [pH 8], 1 mM EDTA, 65 mM KCl, 10% glycerol, 200 μg/ml BSA) containing 6.5 pmol ^{32}P-labeled JG02attDOT substrate (17) and amounts of wild-type and mutant IntDOT protein preparations that shifted one-half of the IHF-*attDOT* complex (see above). The reaction mixtures were incubated at 37°C for 15 min, and the reactions were stopped by mixing the reaction mixtures with an equal volume of 90% formamide in Tris-EDTA. Samples were then boiled for 5 min, run on 15% polyacrylamide TBE buffer-urea gels, and electrophoresed at 400 V for 120 min. Gels were then exposed to phosphorimager screens and visualized using the Fuji FLA-3000 phosphorimager and Fujifilm Image Gauge software (Macintosh v.3.4). The assays were performed at least in triplicate.

RESULTS

Comparative modeling of the IntDOT CB domain. Although IntDOT is a member of the tyrosine family of recombinases, there are no crystal structures available that provide insight into the mechanism of recombination. The CB domains of these proteins interact with DNA in the formation of the higher-order recombination complexes, and recent work on Flp showed that DNA binding specificity is governed largely at the binding step. The CB domains are also involved in protein-protein interactions (2, 29, 30). Thus, amino acid substitutions could potentially affect several steps in the recombination pathway. We chose a homology modeling approach because there are structures available for the CB domains of other tyrosine recombinases that could be used to model IntDOT interactions with DNA. Thus, we used the information to design a set of alanine substitution mutants that could be analyzed for recombination proficiency in vivo and for possible alterations in catalytic activity. If successful, this approach could have significant implications for studies of proteins that have not been analyzed by physical methods.

The IntDOT model (Fig. 1) was derived from the crystal structure of Cre bound covalently to its Holliday junction DNA intermediate (9) and the cocrystal structure of λ Int bound to core-type DNA (1). The Cre and λ Int structures had advantages as modeling templates because they show several interesting higher-order interactions that participate in site-specific recombination, including DNA binding and interprotein contacts. The IntDOT residues involved in these functions can therefore be reasonably extrapolated from an appropriate model built from these template structures. However, because DNA binding sites, accessory factors, and contact surfaces can be expected to evolve independently for different recombi-

nases, the interactions displayed by the model must be validated with subsequent experimental work.

To construct a three-dimensional structural model for IntDOT, several computational analyses were performed. First, protein threading and fold recognition algorithms were applied to identify sequence-structure alignments between IntDOT and other proteins in the Protein Data Bank database. This process yielded significant matches to tyrosine recombinase family members with known structures. Second, inconsistencies in the threading results were identified and evaluated through construction of multiple-sequence alignments. Third, the most consistent sequence-structure alignment, derived from both Cre and λ Int, was used to construct a three-dimensional all-atom structure for IntDOT through homology modeling. During this process the IntDOT residues were formed into a structure closely resembling the known template structures of Cre and λ Int. A molecular dynamics simulation was subsequently applied to relax the structure and minimize the energy of backbone and side chain conformations, and a loop-building process was also applied to predict the best position and orientation for residues in flexible protein regions that are not well conserved.

The levels of identity between IntDOT and the CB domain sequences of other tyrosine recombinases, including Cre, XerD, λ Int, and HP1 integrase (HP1 Int) (12), varied from 10% (Cre) to 22% (HP1 Int) (Fig. 2). Homology modeling typically provides superior results when the level of identity between the query sequence and the structural template is 50% or higher. However, we previously constructed a model of the λ Int CB domain that was later determined to have predicted the CB domain structure when the model was compared with a subsequently published crystal structure. Although λ Int shared only 10.6% identity with the selected Cre template, a useful model of the CB domain could be created because of the strong structural conservation found among several different members of the tyrosine recombinase family (26). There are other examples of structural conservation at similar low levels of sequence identity. For example, we previously showed that the CB domains of Cre and XerD exhibit a conserved three-dimensional fold, despite sharing only 13.5% identical amino acids in this region (26).

IntDOT exhibits a sequence profile that was previously reported for the tyrosine recombinase family CB domain (26). In the present work, the consistency observed among multiple-sequence alignments between IntDOT and the CB domains of either Cre, XerD, HP1 Int, or λ Int indicates that the alignments are generally correct and that IntDOT shares a common CB domain structure with these recombinases. Similar to our previous work with λ Int, the highly consistent sequence alignments also suggest that the overall fold and shape of the IntDOT model are likely to be accurate. However, the low level of identity between IntDOT and the Cre and λ Int template structures (Fig. 2) suggests that the alpha-carbon positions of the polypeptide chain show a root mean square deviation of up to 3 to 4 Å from that predicted in the model.

Although the IntDOT CB domain is only 9.9% identical to the Cre CB domain, there were minimal differences in length, including just two residues inserted at the junction of CB helices 1 and 2, four residues inserted between helices 2 and 3, and one residue inserted between helices 3 and 4 (Fig. 2). The most significant difference observed between IntDOT and Cre was in the region containing helix 5 and the linker that connects the catalytic domain; this segment was nine residues shorter in IntDOT and appears to be more closely related in structure to λ Int in this region. It was reported previously that such loops between secondary structure elements typically diverge among tyrosine recombinase CB domains, presumably due to the reduced functional restrictions on their sequences and their potential roles in independent evolving functions, such as DNA binding (21). The positions and orientations of residues in these weakly conserved regions should therefore be considered approximate. For example, the specific atomic interactions between amino acid side chains and the bound DNA molecule may not be predicted correctly in the model; however, the identities of residues comprising the predicted DNA-binding interface are likely to be correct. In any case, the DNA in the model is not *attDOT*, and sequence-dependent interactions cannot be examined directly.

TABLE 3. Amino acid residues in proximity to DNA

Substitution	Location[a]	Bases[b]	Backbone[b]	Distance (Å)[c]
K129	T1-2	−	+	1–2
K131	T1-2	+	+	1–2
T133	H2	+	+	2–3
L135	H2	+	−	2–3
K136	H2	+	−	1–2
Y137	H2	−	+	2–3
R138	H2	−	+	3–4
T139	H2	+	+	2–3
K142	H2	−	+	2–3
H143	H2	−	+	3–4
H179	T3-4	−	+	3–4
C180	T3-4	−	+	3–4
N183	H4	+	+	1–2
T184	H4	+	−	1–2
W186	H4	+	−	1–2
L187	H4	+	−	1–2
Y188	H4	−	+	2–3
C190	H4	+	−	3–4
R193	H4	−	+	3–4
T194	H4	−	+	3–4

[a] The secondary structure contains the following structures: helix 1 (H1), a turn between helix 1 and helix 2 (T1-2), helix 2 (H2), a turn between helix 2 and helix 3 (T2-3), helix 3 (H3), a turn between helix 3 and helix 4 (T3-4), and helix 4 (H4).

[b] +, hydrogen bond or van der Waals contact with the DNA bases or backbone; −, no hydrogen bond or van der Waals contact with the DNA bases or backbone.

[c] Approximate predicted distance between atoms in the amino acid side chain that are separated from DNA atoms at their closest point.

Through inspection of the model, IntDOT residues in both conserved and nonconserved regions of the CB domain were identified proximal to a bound core-type DNA site (Table 3). The predicted residues were classified according to their proximity to the DNA molecule in the model and the potential for various side chain conformations to interact with either the DNA bases or phosphate backbone. We decided to test experimentally how mutations at these positions affect recombination in vivo and DNA binding, cleavage, and ligation in vitro.

Importance of predicted CB site amino acids for integration in vivo. Site-directed mutagenesis was used to construct alanine substitutions at 15 representative positions in the IntDOT

TABLE 4. Results of the in vivo integration assays and the in vitro DNA binding, ligation, and cleavage assays with the IntDOT mutants

Substitution[a]	Recombination frequency[b]	Gel shift assay (dilution factor)[c]	Ligation reaction (% ligated)[e]	Cleavage reaction (% cleaved)[f]
Wild type	10^{-3}	1/32	60	13
K129A	$<10^{-8}$	ND[d]	ND	ND
K131A	10^{-3}	ND	ND	ND
T133A	5×10^{-5}	ND	ND	ND
L135A	5×10^{-8}	1/16	62	17
K136A	5×10^{-6}	ND	ND	ND
Y137A	$<10^{-8}$	ND	ND	ND
R138A	10^{-6}	ND	ND	ND
T139A	2×10^{-5}	1/16	60	2.5
K142A	6×10^{-7}	1/16	60	6.5
H143A	2×10^{-5}	1/16	15	<1
H179A	3×10^{-6}	1/16	11	5
C180A	5×10^{-7}	ND	ND	ND
N183A	$<10^{-8}$	1/16	42	<1
W186A	3×10^{-3}	ND	ND	ND
T194A	9×10^{-7}	1/16	11	<1

[a] The first letter indicates the amino acid at the position in the wild-type protein. A indicates an alanine substitution.
[b] Integration frequencies were calculated by dividing the number of integrants by the number of recipients.
[c] The gel shift assays were performed in the presence of IHF as described in the text. The dilution factor is the dilution of the IntDOT preparation that shifts approximately one-half of the IHF-*attDOT* DNA complex.
[d] ND, not determined.
[e] Ligation reactions were performed as described in Materials and Methods. The values are the percentages of the substrate ligated in 15 min at 37°C. The amount of IntDOT used for the wild type (1/32 dilution) was one-half the amount used for the mutants (1/16 dilution). All experiments were repeated at least three times.
[f] Cleavage reactions were performed as described in Materials and Methods. The values are the percentages of the substrate cleaved at 37°C. The amount of IntDOT used for the wild type (1/2 dilution) was one-half the amount used for the mutants (undiluted). All experiments were repeated at least three times.

CB domain. Target residues were selected based on their predicted ability to interact with DNA in the model. The majority of target residues were located on helices 2 (T133 to H143) and 4 (N183 to T194), which are predicted to be the orthogonally crossed helices that straddle the major groove of the core DNA site (Fig. 1). Additional substitutions were made in the loop connecting helices 1 and 2 (K129A and K131A) and also in the loop connecting helices 3 and 4 (H179A and C180A).

In order to test IntDOT mutants for integration efficiency, the *Bacteroides* in vivo integration assay was used. In the in vivo system, conjugative mating and transfer of pDJE2.3 containing wild-type or mutant *intDOT* and *ermR* genes are performed using an *E. coli* donor and a *Bacteroides* recipient (5). Because pDJE2.3 cannot replicate in *Bacteroides*, plasmids must integrate into the chromosome by IntDOT-mediated recombination to generate Erm^r colonies (5). Mating assays were repeated three times for each mutant, and the level of integration is reported below as the average frequency of recombination. The results of these experiments are shown in Table 4. Of the 15 substitutions tested, only 2 (K131A and W186A) resulted in wild-type levels of integration (10^{-3} integrant per recipient), while three (K129A, Y137A, and N183A) resulted in undetectable levels of recombination ($<10^{-8}$ integrant per recipient). The remaining 10 mutants displayed recombination frequencies ranging from 10^2- to 10^4-fold less than the wild-type level. The inability of the mutant proteins to promote integration was not due to instability of the proteins. Several strains containing plasmids encoding mutant proteins made levels of IntDOT comparable to the wild-type levels as shown by DNA binding activities of partially purified proteins, as described below.

Binding of mutant IntDOT proteins to *attDOT*. We observed previously that the His_6-tagged version of the wild-type protein was active in our in vivo integration assay. Initially, we constructed amino-terminal His_6-tagged versions of the 15 mutants in the pET28 vector. To assess the ability of mutant IntDOT proteins to bind *attDOT* DNA, we purified the wild-type and 15 mutant proteins and performed gel shift assays. The 320-bp *attDOT* DNA fragment used in this assay was previously shown to be the minimal *attDOT* region. Native IntDOT is unable to retard migration of this *attDOT* DNA unless IHF is also provided (17). We also showed previously that IHF substitutes for the *Bacteroides* host factor in vitro (7). Presumably, IHF binds nonspecifically to *attDOT* DNA and promotes intasome formation with IntDOT through stabilization of bent DNA conformations. When wild-type IntDOT was added to a reaction mixture containing this complex, a new, more slowly migrating band containing IHF and IntDOT was detected (data not shown). Because formation of this complex requires IHF, the complex could be similar to the complex formed in λ *attL*, where IHF bends DNA, allowing an Int monomer to simultaneously interact with an arm site and a core site via different DNA binding domains (15, 16).

Our results revealed that five of the His_6-tagged mutants (K129A, T133A, K136A, C180A, and T194A) showed altered or defective activity in the gel shift assay. These mutants were not studied further. Ten of the His_6-tagged mutant proteins (L131A, L135A, L137A, R138A, T139A, K142A, H143A, H179A, N183A, and T194A) were able to form a complex similar to the His_6-tagged wild-type protein (data not shown). However, we found that the His_6-tagged wild-type protein was much less active than a partially purified native wild-type protein in our in vitro recombination system (18). Thus, we made pET28 versions (lacking the His_6 tag) of the wild-type protein for seven of the remaining mutants (L135A, T139A, K142A, H143A, H179A, N183A, and T194A) and partially purified the proteins as described in Materials and Methods.

The wild-type and seven mutant proteins were analyzed using the gel shift assay. The amount of the wild-type IntDOT preparation required to shift one-half of the IHF-*attDOT* complex was approximately one-half the amount of the mutant preparations required (Fig. 3, Table 4, and data not shown). These results indicated that the mutant proteins were stable and could be isolated from cells at concentrations comparable to those of the wild-type IntDOT protein. In the cleavage and ligation experiments reported below, we used amounts of wild-type and mutant proteins that shift one-half of the IHF-*attDOT* complex. However, because the *attDOT* fragment used in the assay contains core-type and arm-type sequences, we cannot exclude the possibility that the seven mutant proteins may have subtle defects in DNA binding that are not apparent in the assay used. Additionally, cooperative interactions between IntDOT monomers could mask subtle defects in DNA binding.

Ability of IntDOT mutants to ligate and cleave *attDOT* DNA. The ligation assay measures the ability of a protein to catalyze the ligation of a nick containing a 5′ hydroxyl with an activated

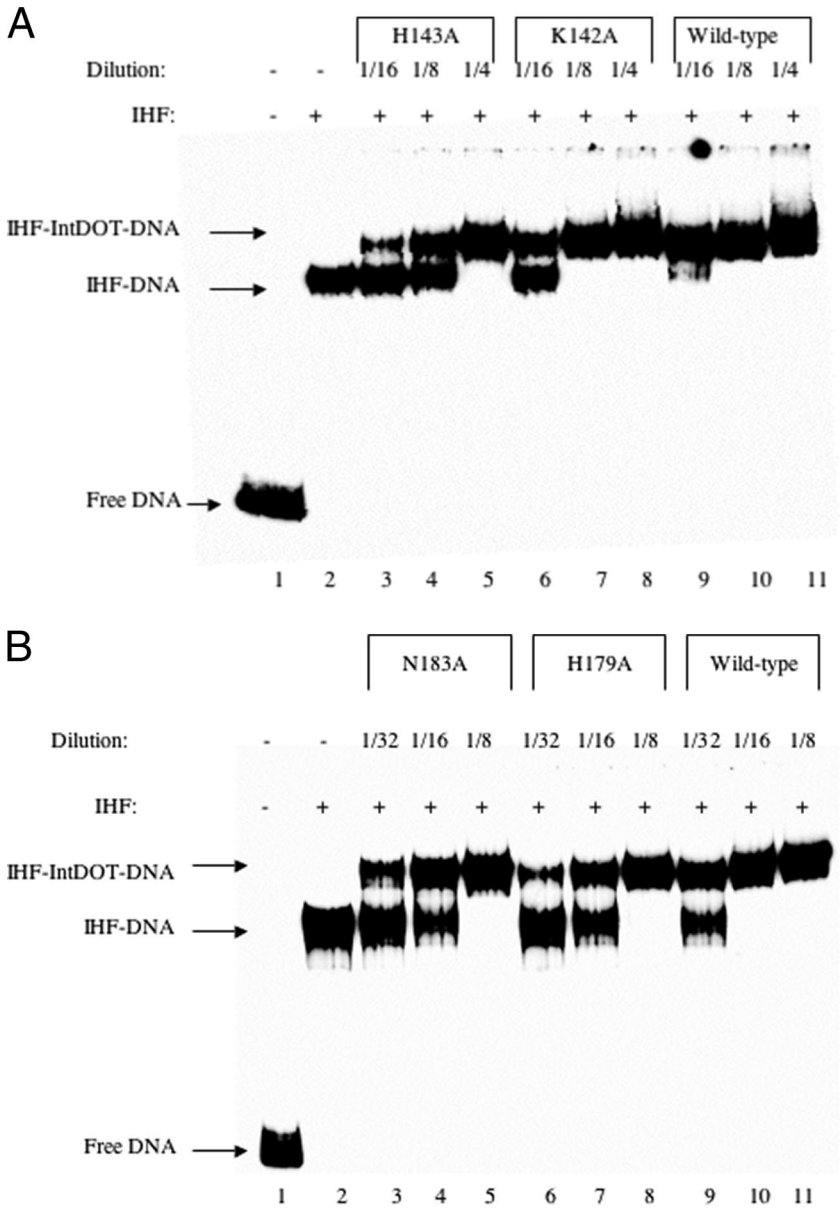

FIG. 3. Electrophoretic mobility shift assays of wild-type and mutant IntDOT proteins bound to *attDOT* DNA. The experimental details are described in Materials and Methods. (A) Lane 1, *attDOT* DNA; lane 2, *attDOT* DNA and IHF; lanes 3, 4, and 5, *attDOT* DNA, IHF, and the H143A protein diluted 1/16, 1/8, and 1/4, respectively; lanes 6, 7, and 8, *attDOT* DNA, IHF, and the K142A protein diluted 1/16, 1/8, and 1/4, respectively; lanes 9, 10, and 11, *attDOT* DNA, IHF, and the wild-type protein diluted 1/16, 1/8, and 1/4, respectively. (B) Lane 1, *attDOT* DNA; lane 2, *attDOT* DNA and IHF; lanes 3, 4, and 5, *attDOT* DNA, IHF, and the N183A protein diluted 1/32, 1/16, and 1/8, respectively; lanes 6, 7, and 8, *attDOT* DNA, IHF, and the H179A protein diluted 1/32, 1/16, and 1/8, respectively; lanes 9, 10, and 11, *attDOT* DNA, IHF, and the wild-type protein diluted 1/32, 1/16, and 1/8, respectively.

substrate containing a 3′ pNP group. The pNP mimics the 3′-phosphotyrosine intermediate of the recombination reaction. The ligation assay was performed by incubating the wild-type and mutant proteins with the radiolabeled 3′ pNP-activated *attDOT* substrate. We analyzed the same seven mutants (L135A, T139A, K142A, H143A, H179A, N183A, and T194A) that had defective integration activity in vivo but still shifted *attDOT* DNA in vitro for the ability to perform the ligation reaction. Figure 4 shows representative gels for assays performed with wild-type and several mutant IntDOT proteins. All seven mutants were able to perform the ligation reaction and form a phosphodiester bond, although some were more efficient than others.

The mutants were also tested for cleavage activity in vitro. Cleavage of the DNA substrate is necessary to form the 3′-phosphotyrosine intermediate of the recombination reaction. This assay utilized an *attDOT* substrate with the top strand (*attDOT-T4*) ending 3 bp 3′ of the site of cleavage and strand exchange (17). If the protein is able to cleave the substrate, the 3-bp fragment should diffuse away and leave the protein

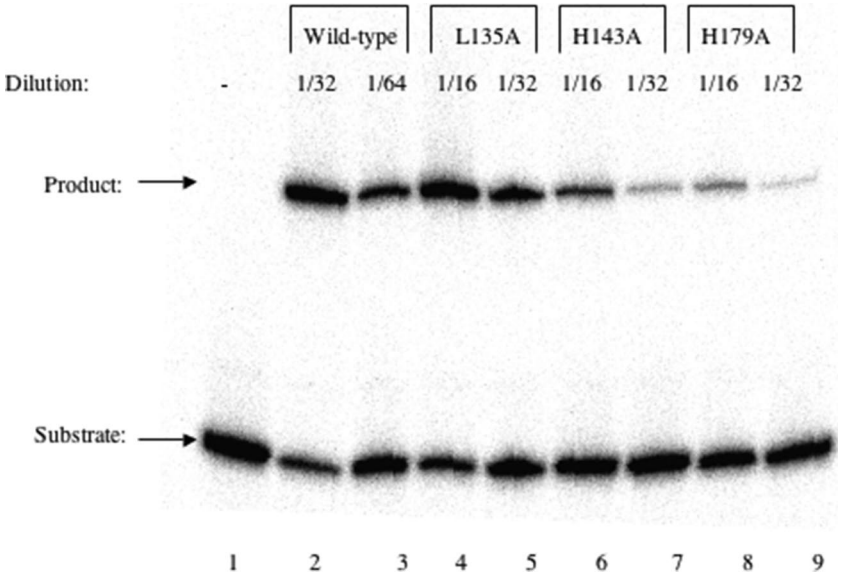

FIG. 4. Ligation activities of wild-type and mutant proteins. The ligation assays were performed and analyzed as described in Materials and Methods. Lane 1, DNA substrate in the absence of protein; lanes 2 and 3, wild-type protein diluted 1/32 and 1/16, respectively; lanes 4 and 5, L135A protein diluted 1/16 and 1/32, respectively; lanes 6 and 7, H143A protein diluted 1/16 and 1/32, respectively; lanes 8 and 9, H179A protein diluted 1/16 and 1/32, respectively.

covalently bound to the substrate. This complex is detected as a more slowly migrating band on a 4 to 20% Tris-glycine gradient gel. The results of cleavage assays for the wild-type protein and several of the mutant proteins are shown in Fig. 5. To prove that the products of these reactions were, in fact, covalently bound protein-DNA intermediates, proteinase K was added to some samples (17). The more slowly migrating product band disappeared, indicating that the complexes contained covalent protein-DNA linkages (data not shown).

The combined results of the ligation and cleavage assays, as well as the in vivo integration assays, are shown in Table 4. An intriguing finding is that many of the mutants are proficient in the ligation reaction but defective in the cleavage reaction. The L135A, T139A, K142A, and N183A proteins were the most efficient proteins for performing the ligation reaction and appear to be comparable to the wild-type protein. The H143A, H179A, and T194A proteins were deficient in the ability to perform the ligation reaction. Of the seven mutants tested, only the L135A and K142A mutants were able to perform the cleavage step at a level comparable to that of the wild-type protein. The remaining mutants displayed severely depressed levels of cleavage activity.

DISCUSSION

Homology modeling can predict various aspects of protein structure and function in the absence of other structural information. Although the quality of predicted structures is improving rapidly as knowledge of protein biochemistry increases and additional protein structures are reported in the literature, it is prudent to demonstrate the biological relevance and validity of a model by correlating its predictions with experimental results obtained in the laboratory.

In this report, we used homology modeling to predict the

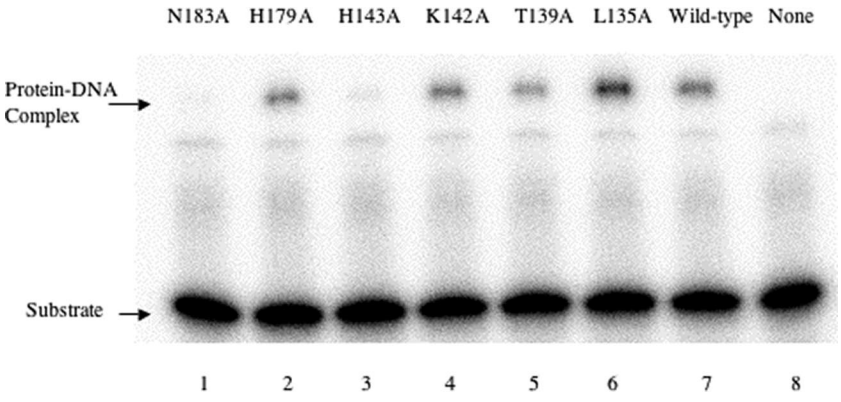

FIG. 5. DNA cleavage assays of wild-type and mutant proteins. The cleavage reactions were performed as described in Materials and Methods. The proteins used in the assays are indicated above the gel. Lane 8 contained a reaction mixture with no added protein.

structure of the IntDOT CB domain. Sequence-structure alignments with Cre, XerD, HP1 Int, and λ Int indicated that the CB domain of IntDOT shares the common protein fold reported for other members of the tyrosine recombinase family. Because the IntDOT model was derived from the Cre cocrystal structure (9, 10) showing interactions with a bound core-type site, the IntDOT model could be examined in the context of the high-order interactions shown in the parent structure. The residues that we mutated were predicted to interact with the bound DNA molecule and reside on helices 2 (T133 to H143) and 4 (N183 to T194) and in the loops connecting helices 1 and 2 (K129 and K131) and helices 3 and 4 (H79 and C180).

Thirteen of 15 mutant IntDOT proteins showed a significant decrease in recombination activity in vivo (Table 4), demonstrating the functional importance of the residues predicted by the model to interact with core-type DNA. The three substituted positions that showed a complete loss of recombination activity are located on the loop between helices 1 and 2 (K129A), in the middle of helix 2 (Y137A), and at the beginning of helix 4 (N183A). All eight substitutions in helix 2 resulted in at least a 10^2-fold decrease in recombination compared to the wild-type IntDOT, further supporting a role of this region in binding to DNA.

Additional assays were performed to determine at which step in the recombination reaction (i.e., DNA binding, cleavage, or ligation) the mutant proteins are deficient. Gel mobility shift assays with partially purified wild-type and mutant proteins and a minimal *attDOT* DNA site showed that several IntDOT mutants (L135A, T139A, K142A, H143A, H179A, N183A, and T194A) were defective for recombination in vivo but retained the ability to form IHF-dependent complexes with *attDOT* DNA (Table 4). It appears that these proteins can form a higher-order complex with Int and *attDOT* DNA in a manner similar to that of the wild-type protein. However, because we have not yet characterized the protein composition of the IHF-dependent complex, we cannot rule out the possibility that cooperative interactions mask defects of the mutant proteins. In addition, it is possible that there are subtle differences that are not detected by the gel shift assay.

We also tested these mutants for the ability to cleave and ligate DNA in vitro (Table 4) and identified three residues that may play a role in the catalytic activity of IntDOT. The T139A, H143A, N183A, and T194A proteins are particularly interesting because they are defective in DNA cleavage, yet retain substantial ligation activity. With the exception of mutants that contain substitutions of the conserved catalytic tyrosine residues, this phenotype has not been reported previously for any other mutant tyrosine recombinase. The H143 residue is predicted to interact with DNA approximately three bases from the cleavage site on the cleaved strand, and the T139 residue interacts with the cleaved strand near H143 (Fig. 6). The T194 residue is predicted to interact with a DNA base partner in the uncleaved strand near the site of cleavage. The location of N183 in the model is striking because N183 is predicted to interact with the DNA adjacent to the cleavage site. The ligation proficiencies of the H143A, N183A, and T194A proteins indicate that these residues are required only during DNA cleavage. We suggest that the IntDOT protein conformation may differ during strand cleavage and ligation in such a way

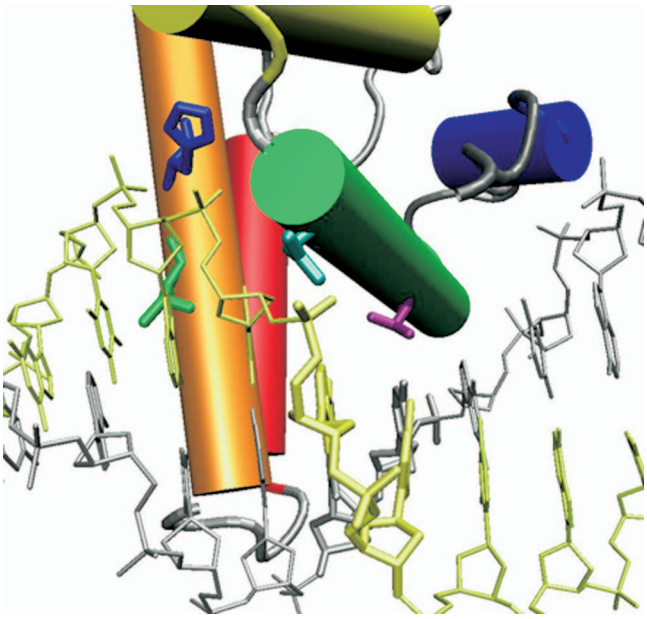

FIG. 6. Locations of amino acid positions yielding a cleavage-specific defect in the model. The DNA bases and backbone flanking the predicted cleaved phosphodiester bond are indicated by a slightly larger diameter. N183 (cyan), H143 (blue), and T139 (green) all interact with the cleaved DNA strand (yellow). T194 (purple) is proximal to the uncleaved DNA strand (gray). N183 interacts with the DNA adjacent to the predicted point of DNA cleavage (between the bold yellow bases).

that N183 and possibly H143 and T194 are required to position the DNA only when it is in the cleavage-activated conformation. An alternative interpretation is that the alanine substitutions at positions 143, 183, and 194 negatively affect strand cleavage; however, the mechanism by which alanine would interfere with DNA binding is not apparent from the model and thus more difficult to explain.

When examined in our previous CB domain alignment (26), N183 is not a conserved residue in the tyrosine recombinase family but is located adjacent to one of the most conserved polar residues in the alignment, T184 (which corresponds to λ Int S139). Because the N183 position is not conserved and is located in a highly variable region of the CB domain alignment, N183 may have evolved to play a unique role in the unusual recombination mechanism catalyzed by IntDOT. Additional work is planned to further investigate the role of N183.

The importance of helices 2 and 4 and their adjacent sequences in the IntDOT CB domain is confirmed by the importance of these structural elements in other recombinases. In the various crystal structure models of a Cre dimer bound to *loxP* DNA (8–10, 32), helices 2 and 4 of Cre are orthogonally crossed, and residues on these two helices and their adjacent turns penetrate the major DNA groove and make contact with the DNA backbone and bases (10). We previously found that CB domain helices 2 and 4 are conserved among many tyrosine recombinases and they apparently provide an important structural framework for binding to core-type DNA sites that is widely used in this protein family (26). Genetic and biochemical experiments have also indicated that the analogous α-helices in λ Int and HK022 Int are important for the specificity

and affinity of binding to core-type DNA sites (6, 11, 25, 28), Studies of Flp-DNA interactions indicate that although DNA binding specificity is mediated largely at the binding step, interactions of the protein with DNA can also affect catalysis (30).

In conclusion, the homology model of IntDOT bound to DNA allowed us to predict amino acids in the CB domain that are important for interactions with DNA. The experimental analysis of the model presented here showed that the model provides a useful tool to investigate the structure and function of IntDOT. The results indicate that some residues in the CB domain of IntDOT not only interact with DNA but may also contribute to catalysis by positioning DNA during the DNA cleavage reaction.

ACKNOWLEDGMENTS

We thank M. McCurdy, J. Laprise, R. Jeters, and S. Yoneji for technical assistance and help preparing the manuscript. We also thank members of the Gardner and Salyers laboratories for helpful discussions and advice.

This work was supported by National Institutes of Health grants GM28717 and AI22386.

REFERENCES

1. **Aihara, H., H. J. Kwon, S. E. Nunes-Duby, A. Landy, and T. Ellenberger.** 2003. A conformational switch controls the DNA cleavage activity of lambda integrase. Mol. Cell **12:**187–198.
2. **Azaro, M. A., and A. Landy.** 2002. λ integrase and the lambda Int family, p. 119–148. *In* N. L. Craig, R. Craigie, M. Gellert, and A. M. Lambowitz (ed.), Mobile DNA II. ASM Press, Washington, DC.
3. **Bauer, C. E., S. D. Hesse, R. I. Gumport, and J. F. Gardner.** 1986. Mutational analysis of integrase arm-type binding sites of bacteriophage lambda. Integration and excision involve distinct interactions of integrase with arm-type sites. J. Mol. Biol. **192:**513–527.
4. **Bushman, W., J. F. Thompson, L. Vargas, and A. Landy.** 1985. Control of directionality in lambda site specific recombination. Science **230:**906–911.
5. **Cheng, Q., B. J. Paszkiet, N. B. Shoemaker, J. F. Gardner, and A. A. Salyers.** 2000. Integration and excision of a *Bacteroides* conjugative transposon, CTnDOT. J. Bacteriol. **182:**4035–4043.
6. **Cheng, Q., B. M. Swalla, M. Beck, R. Alcaraz, Jr., R. I. Gumport, and J. F. Gardner.** 2000. Specificity determinants for bacteriophage Hong Kong 022 integrase: analysis of mutants with relaxed core-binding specificities. Mol. Microbiol. **36:**424–436.
7. **Cheng, Q., N. Wesslund, N. B. Shoemaker, A. A. Salyers, and J. F. Gardner.** 2002. Development of an in vitro integration assay for the *Bacteroides* conjugative transposon CTnDOT. J. Bacteriol. **184:**4829–4837.
8. **Gopaul, D. N., F. Guo, and G. D. Van Duyne.** 1998. Structure of the Holliday junction intermediate in Cre-loxP site-specific recombination. EMBO J. **17:** 4175–4187.
9. **Guo, F., D. N. Gopaul, and G. D. Van Duyne.** 1999. Asymmetric DNA bending in the Cre-loxP site-specific recombination synapse. Proc. Natl. Acad. Sci. USA **96:**7143–7148.
10. **Guo, F., D. N. Gopaul, and G. D. van Duyne.** 1997. Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. Nature **389:**40–46.
11. **Han, Y. W., R. I. Gumport, and J. F. Gardner.** 1994. Mapping the functional domains of bacteriophage lambda integrase protein. J. Mol. Biol. **235:**908–925.
12. **Hickman, A. B., S. Waninger, J. J. Scocca, and F. Dyda.** 1997. Molecular organization in site-specific recombination: the catalytic domain of bacteriophage HP1 integrase at 2.7 A resolution. Cell **89:**227–237.
13. **Jones, D. T.** 1999. GenTHREADER: an efficient and reliable protein fold recognition method for genomic sequences. J. Mol. Biol. **287:**797–815.
14. **Jones, D. T.** 1999. Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. **292:**195–202.
15. **Kim, S., L. Moitoso de Vargas, S. E. Nunes-Duby, and A. Landy.** 1990. Mapping of a higher order protein-DNA complex: two kinds of long-range interactions in lambda *attL*. Cell **63:**773–781.
16. **MacWilliams, M. P., R. I. Gumport, and J. F. Gardner.** 1997. Mutational analysis of protein binding sites involved in formation of the bacteriophage lambda *attL* complex. J. Bacteriol. **179:**1059–1067.
17. **Malanowska, K., A. A. Salyers, and J. F. Gardner.** 2006. Characterization of a conjugative transposon integrase, IntDOT. Mol. Microbiol. **60:**1228–1240.
18. **Malanowska, K., S. Yoneji, A. A. Salyers, and J. F. Gardner.** 2007. CTnDOT integrase performs ordered homology-dependent and homology-independent strand exchanges. Nucleic Acids Res. **35:**5861–5873.
19. **McGuffin, L. J., and D. T. Jones.** 2003. Improvement of the GenTHREADER method for genomic fold recognition. Bioinformatics **19:**874–881.
20. **Nash, H. A.** 1996. The Hu and IHF proteins: accessory factors for complex protein-DNA assemblies, p. 149–179. *In* E. C. C. Lin and A. S. Lynch (ed.), Regulation of gene expression in Escherichia coli. RG Landes, Austin, TX.
21. **Nunes-Duby, S. E., H. J. Kwon, R. S. Tirumalai, T. Ellenberger, and A. Landy.** 1998. Similarities and differences among 105 members of the Int family of site-specific recombinases. Nucleic Acids Res. **26:**391–406.
22. **Schwede, T., J. Kopp, N. Guex, and M. C. Peitsch.** 2003. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. **31:**3381–3385.
23. **Studier, F. W., A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff.** 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. **185:**60–89.
24. **Subramanya, H. S., L. K. Arciszewska, R. A. Baker, L. E. Bird, D. J. Sherratt, and D. B. Wigley.** 1997. Crystal structure of the site-specific recombinase, XerD. EMBO J. **16:**5178–5187.
25. **Swalla, B. M., E. H. Cho, R. I. Gumport, and J. F. Gardner.** 2003. The molecular basis of co-operative DNA binding between lambda integrase and excisionase. Mol. Microbiol. **50:**89–99.
26. **Swalla, B. M., R. I. Gumport, and J. F. Gardner.** 2003. Conservation of structure and function among tyrosine recombinases: homology-based modeling of the lambda integrase core-binding domain. Nucleic Acids Res. **31:**805–818.
27. **Thompson, J. F., U. K. Snyder, and A. Landy.** 1988. Helical-repeat dependence of integrative recombination of bacteriophage lambda: role of the P1 and H1 protein binding sites. Proc. Natl. Acad. Sci. USA **85:**6323–6327.
28. **Tirumalai, R. S., H. J. Kwon, E. H. Cardente, T. Ellenberger, and A. Landy.** 1998. Recognition of core-type DNA sites by lambda integrase. J. Mol. Biol. **279:**513–527.
29. **Van Duyne, G. D.** 2002. A structural view of tyrosine recombinase site-specific recombination. ASM Press, Washington, DC.
30. **Whiteson, K. L., and P. A. Rice.** 2008. Binding and catalytic contributions to site recognition by flp recombinase. J. Biol. Chem. **283:**11414–11423.
31. **Woodfield, G., C. Cheng, S. Shuman, and A. B. Burgin.** 2000. Vaccinia topoisomerase and Cre recombinase catalyze direct ligation of activated DNA substrates containing a 3′-para-nitrophenyl phosphate ester. Nucleic Acids Res. **28:**3323–3331.
32. **Woods, K. C., S. S. Martin, V. C. Chu, and E. P. Baldwin.** 2001. Quasi-equivalence in site-specific recombinase structure and function: crystal structure and activity of trimeric Cre recombinase bound to a three-way Lox DNA junction. J. Mol. Biol. **313:**49–69.