Proc. Natl. Acad. Sci. USA
Vol. 94, pp. 6104–6109, June 1997
Biochemistry

The catalytic domain of λ site-specific recombinase

RADHAKRISHNA S. TIRUMALAI, ERICA HEALEY, AND ARTHUR LANDY*

Division of Biology and Medicine, Brown University, Providence, RI 02912

Communicated by Allan Campbell, Stanford University, Stanford, CA, April 9, 1997 (received for review January 7, 1997)

ABSTRACT **The *Escherichia coli* phage λ integrase protein (Int) belongs to the large Int family of site-specific recombinases. It is a heterobivalent DNA binding protein that makes use of a high energy covalent phosphotyrosine intermediate to catalyze integrative and excisive recombination at specific chromosomal sites (*att* sites). A 293-amino acid carboxy-terminal fragment of Int (C65) has been cloned, characterized, and used to further dissect the protein. From this we have cloned and characterized a 188-amino acid, protease-resistant, carboxy-terminal fragment (C170) that we believe is the minimal catalytically competent domain of Int. C170 has topoisomerase activity and converts *att* suicide substrates to the covalent phosphotyrosine complexes characteristic of recombination intermediates. However, it does not show efficient binding to *att* site DNA in a native gel shift assay. We propose that λ Int consists of three functional and structural domains: residues 1–64 specify recognition of "arm-type" DNA sequences distant from the region of strand exchange; residues 65–169 contribute to specific recognition of "core-type" sequences at the sites of strand exchange and possibly to protein–protein interactions; and residues 170–356 carry out the chemistry of DNA cleavage and ligation. The finding that the active site nucleophile Tyr-342 is in a uniquely protease-sensitive region complements and reinforces the recently solved C170 crystal structure, which places Tyr-342 at the center of a 17-amino acid flexible loop. It is proposed that C170 is likely to represent a generic Int family domain that thus affords a specific route to studying the chemistry of DNA cleavage and ligation in these recombinases.**

The *Escherichia coli* phage λ integrase protein (Int) belongs to a large family of site-specific DNA recombinases that execute rearrangements between DNA sequences with little or no sequence homology (for reviews see refs. 1–3). This family was initially defined functionally and by the strict conservation of four residues critical for catalysis: Arg-212-His-308-Arg-311-Tyr-342 (λ Int numbering) (4, 5). The recent solution of the x-ray crystal structure of the catalytic domain of λ Int suggests an even higher degree of conservation among the members of the Int family: that they share a common protein fold for the domain that executes the chemistry of DNA cleavage and ligation (6).

As first proposed by Campbell (7) λ Int catalyzes integrative and excisive recombination between pairs of specific DNA target sites, *att*P/*att*B and *att*L/*att*R, respectively. The core of each *att* sequence consists of two Int binding sites, one at each strand cleavage site, separated by a 7-bp "overlap region." A reciprocal exchange of the "top" strands at the left boundary of the overlap region generates a Holliday junction recombination intermediate that is then resolved by exchange of the "bottom" strands at the right boundary of the overlap region. During DNA cleavage, a tyrosine hydroxyl attacks the scissile phosphate, nicking the DNA and forming a 3′ phosphotyrosine-linked DNA complex. This covalent protein–DNA intermediate is resolved when the 5′-terminal hydroxyl of the invading DNA strand attacks the phosphotyrosine linkage and displaces the protein. Additional binding sites for Int and accessory proteins are present in the distal arms flanking the core sequences but they are not essential to the chemistry and mechanisms of DNA strand transfer.

The 40-kDa Int protein, which was first purified by Kikuchi and Nash (8), is a type I topoisomerase and a heterobivalent DNA binding protein. Limited proteolysis of Int with chymotrypsin generated two peptides (9). The amino-terminal peptide (residues 1–64) binds with high affinity to the five "arm-type" sites distal of the region of strand exchange. The carboxy-terminal peptide (residues 65–356) binds to the "core-type" sites at the positions of strand cleavages (flanking the 7-bp overlap region) and also has topoisomerase activity. This report describes the identification, isolation, and characterization of what we propose is the minimal domain competent for the DNA cleavage and ligation chemistry of λ site-specific recombination. These results and the observed protease sensitivity of the Tyr-342 active site nucleophile are discussed in relation to the x-ray crystal structure of this domain (6) and the conservation of structure and mechanism among the Int family recombinases.

MATERIALS AND METHODS

Proteins. Bacteriophage λ Int protein was produced from an expression plasmid under the control of a T7 promoter in *E. coli* BL21. The carboxy-terminal fragment of the Int protein encompassing residues 65–356 (called C65) was generated by exonuclease III digestion (10) of the λ Int gene (Istvan Papp, unpublished data) and was introduced into the expression plasmid behind an initiator methionine under the control of a T7 promoter (11). The Y342F mutant of C65 (C65F) was made by swapping the *Bam*HI–*Hin*dIII fragment from λ Int Y342F (λ Int F). The carboxy- terminal fragment comprising residues 170–356 (C170) and its mutant, C170,Y342F (C170F), were both made by PCR using the λ Int gene/λ Int F gene, respectively, as the template. The promoter–proximal primer had a unique *Nde*I restriction site at the region coding for amino acid 170, and the distal primer had a unique *Hin*dIII restriction site downstream of the Int gene. The PCR product was digested with *Nde*I and *Hin*dIII and inserted into the expression plasmid at the corresponding *Nde*I–*Hin*dIII restriction sites. All of the proteins were overexpressed and purified to near homogeneity by a modification of the procedure of Kikuchi and Nash (8) (to be described elsewhere).

Oligonucleotides. Synthetic oligos (HPLC-purified) were obtained from Operon technologies, Alameda, CA and were labeled at their 5′ OH terminus with ^{32}P γ-ATP, using T4 polynucleotide kinase (12). Sequences of the top strand (5′→3′) for each of the double-stranded oligos used in this work are as follows (in each case the complementary strand has a 5′ extension indicated in parenthesis 5′→3′, and the Int core-type recognition sequence is noted as bold caps): (*i*)

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: Int, integrase.

*To whom reprint requests should be addressed.

30-bp core-type DNA for gel-mobility shift: actgctccttttataatgc**CAACTT**tt**T**t(c); (*ii*) 30-bp competitor DNA for gel-mobility shift: actgctccttttataatgctaatccttttt(c); (*iii*) 13-bp core-type DNA for UV-crosslinking: g**c**CAACTTag**T**t(t); and (*iv*) half-*att* site (suicide substrate): aagctgaagatcttctcgag**CAGCTT**-tc**T**a(gttc).

Limited Proteolysis. Protein (25 μg; λ Int, C65, or C170) was limit digested with 1 μg of Arg-C or kallikrein or 50 ng of chymotrypsin or trypsin in a reaction mixture containing 20 mM Tris·HCl (pH 7.5), 10% (vol/vol) glycerol, 300 mM NaCl, 1 mM EDTA, and 1 mM DTT at 25°C. At 0, 30, and 60 min, aliquots were withdrawn and analyzed by 18% SDS/PAGE (13).

RESULTS

Domain Structure. Individual protein domains can often be identified and isolated by limited proteolysis under native conditions because the flexible unstructured interdomain linkers are considerably more susceptible to cleavage (14, 15). Using this approach, it had been shown that limited proteolysis of the bacteriophage λ-Int protein generates two putative domain fragments (9). The smaller amino-terminal fragment (residues 1–64) recognizes and binds to arm-type DNA. The larger carboxy-terminal domain (residues 65–356) binds to core-type DNA and has topoisomerase activity. However, neither of the fragments can catalyze *in vitro* recombination. The carboxy-terminal domain (C65) was cloned, overexpressed, and purified to near homogeneity (see *Materials and Methods*). A derivative in which the catalytic Tyr-342 is converted to Phe (referred to as C65F) also was purified to near homogeneity.

To further dissect the domain structure of Int, purified C65 was subjected to limited proteolysis under native conditions with endoproteinase Arg-C or kallikrein. Gel electrophoresis showed that both proteases yielded specific fragments of approximately the same sizes, 20, 19, and 11 kDa, respectively (Fig. 2*A*). The structural significance of these proteolysis products was suggested by the fact that proteolysis of intact Int with Arg-C or kallikrein yielded the same profiles as observed with C65 (Chris Pargellis, unpublished work). Arg-C and kallikrein are both arginine-specific proteases, so two additional protease specificities also were examined: trypsin, which cleaves after lysine as well as arginine, and chymotrypsin, which cleaves after hydrophobic residues. Both of these proteases yielded digestion products of approximately the same size as the Arg-C and kallikrein for both full length Int and C65 (data not shown). This similarity of digestion products with a range of proteases suggests that the sensitive target sites are not due to unusual properties of selected amino acid residues but probably reflect accessible and/or flexible regions in the global protein structure. We shall first consider the 20- and 11-kDa fragments, and the significance of the 19-kDa fragment protease product will be discussed further below. The 20-kDa fragment was transferred from the gel by electroblotting onto an Immobilon (Millipore) membrane and was subjected to 10 cycles of N-terminal sequencing. The results showed unambiguously that the Arg-C cleavage generating this fragment occurred between Arg-169 and Ala-170. The product of kallikrein cleavage also was sequenced and found to have the same N-terminal sequence. The 11-kDa fragment from an Arg-C digest of C65 also was recovered from an SDS/polyacrylamide gel by electroblotting. Ten cycles of N-terminal sequencing showed that this fragment has the same N terminus as C65. We conclude that the 11- and 20-kDa fragments comprise the upstream and downstream products, respectively, of one, or several closely spaced, Arg-C/kallikrein cleavages of C65. Based on these results, the region from Ala-170 to the carboxyl terminus of Int was amplified by PCR using primers that both added a Met to the N terminus and enabled cloning into the pT7–7 expression vector. The C170 protein was overexpressed and purified to a single band on an overloaded SDS gel stained with Coomassie blue (not shown).

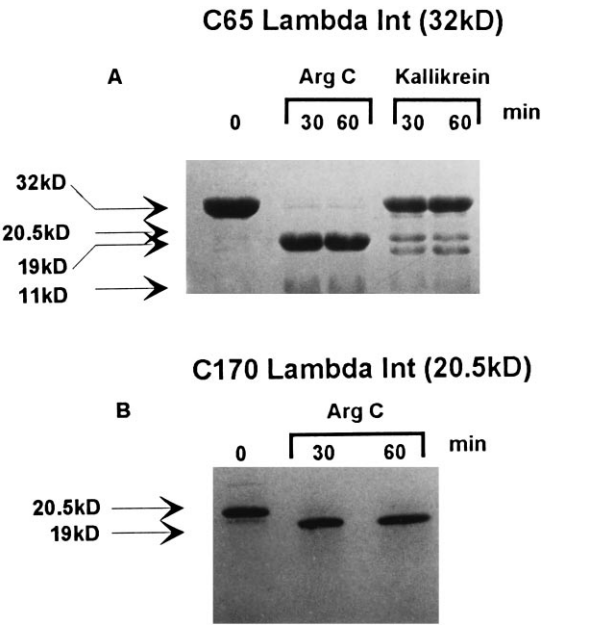

FIG. 2. Limited proteolysis of the carboxy-terminal fragments of λ Int. The carboxy-terminal fragments C65 (amino acids 65–356) (*A*) or C170 (amino acids 170–356) (*B*) were subjected to limited proteolysis with endoproteinase Arg-C or kallikrein (see *Materials and Methods*). At the indicated times, aliquots were withdrawn, the products were analyzed by SDS/PAGE, and the gels were stained with Coomassie blue (see *Materials and Methods*). The approximate sizes of the fragments (*Left*) were determined from their electrophoretic mobility relative to known standards (not shown).

Topoisomerase Activity. As noted above, Int protein is a type I topoisomerase. However, in contrast to its DNA sequence specificity in catalyzing site-specific recombination, it is able to relax supercoiled pBR322 DNA with the same efficiency as plasmid carrying λ *att* site sequences (16). The

```
1
MGRRRSHERR DLPPNLYIRN NGYYCYRDPR TGKEFGLGRD RRIAITEAIQ ANIELFSGHK
61  ↓
HKPLTARINS DNSVTLHSWL DRYEKILASR GIKQKTLINY MSKIKAIRRG LPDAPLEDIT
121                                                 ⇩
TKEIAAMLSG YIDEGKAASA KLIRSTLSDA FREAIAEGHI TTNPVAATRA AKSEVRRSRL
181                    •
TADEYLKIYQ AAESSPCWLR LAMELAVVTG QRVGDLCEMK WSDIVDGYLY VEQSKTGVKI
241
AIPTTLHVDA LGISMKETLD KCKEILGGET IIASTRREPL SSGTVSRYFM RARKASGLSF
301     •   •                                  ⇩  ✱              356
EGDPPTFHEL RSLSARLYEK QISDKFAQHL LGHKSDTMAS QYRDDRGREW DKIEIK
```

FIG. 1. Protease-sensitive sites in λ Int. The sites of cleavage under conditions of limited proteolysis are shown for chymotrypsin (filled arrow) and Arg-C and kallikrein (open arrows). The three residues of the catalytic triad, Arg-212, His-308, and Arg-311, are marked by filled circles, and the active site nucleophile, Tyr-342 is marked by an asterisk.

previously reported finding that topoisomerase function is retained in a proteolytic fragment of Int lacking 64 amino acids from the N terminus was subject to the reservation that the activity might have been due to a small amount of contaminating intact Int (9). We have now eliminated this reservation by showing that the topoisomerase activity of the cloned and purified C65 fragment is indistinguishable from that of full length Int (Fig. 3 *A* and *B*). An agarose gel electrophoresis system in which supercoiled DNA migrates faster than its relaxed topoisomers was used to assay the ability to relax supercoiled pBR322 DNA (Fig. 3).

The same assay was used to characterize the topoisomerase activity of purified C170. The C170 topoisomerase activity was ≈10% of the C65 activity. However, in the presence of 10% dimethyl sulfoxide, the C170 topoisomerase is stimulated ≈4-fold whereas there is no effect on the activity of C65 or full length Int. Thus, under these reaction conditions, C170 (Fig. 3*C*) is ≈40% of that of the two larger proteins (Fig. 3 *A* and *B*). Although the mechanism of DMSO stimulation is not clear, DMSO is known to reduce the stringency of protein–DNA interactions (see *Discussion*).

Binding to Core-Type DNA. To study the binding of C65 and C170 to *att* site DNA, mutant proteins were used in which the active site Tyr-342 nucleophile was changed to Phe, thus eliminating the possibility of forming covalent protein DNA intermediates (see below). A ^{32}P-labeled, 30-bp core-type DNA fragment containing the overlap region and a single core-type Int binding site was incubated with protein and then electrophoresed on a nondenaturing polyacrylamide gel. With C65F, a single retarded band was observed at low protein concentrations, and a second slower moving complex was observed at higher concentrations (Fig. 4*A*). A gel electrophoresis method was used to determine the size of the retarded complexes (17). The faster migrating complex has an approximate size corresponding to one C65 protomer plus one 30-bp oligomer. The slower protein-dependent band had an approximate size corresponding to two C65 protomers complexed with either one 30-bp oligomer or one C65 protomer complexed with two 30 bp oligomers (David Woo, unpublished results). When a smaller core-type fragment of 13 bp was used, only a single band corresponding to a 1:1 molar ratio of protein to DNA was observed (data not shown). A more detailed characterization of the slower moving band is currently in progress, but, for this report, we are primarily interested in obtaining a reference for the C170 binding.

In contrast to C65, the C170 protein did not show any binding to the 30-mer *att* site fragment under these conditions (Fig. 4*B*). Increasing the C170 concentration 5-fold did not result in detectable binding, and increasing it 100- to 200-fold resulted in a smeared pattern of DNA retardation (not shown). To examine the specificity of C65 binding to the *att* site core-type DNA fragment, the binding was carried out in the presence of different concentrations of unlabeled core-type (homologous competitor), or unlabeled non core-type DNA (heterologous competitor), which has an altered consensus binding site, but same base composition. Very little reduction in C65 binding was observed over a range of equimolar to five-fold excess of heterologous competitor, while, the homologous competitor, under similar conditions, reduced the binding significantly (Fig. 4*C*). In contrast to this, even the retarded smear observed with 100–200-fold excess C170 was reduced in direct proportion to the amount of cold competitor (not shown). This suggests that both the affinity and specifity of C170 binding to DNA are greatly reduced. These results suggest that the region of λ Int between residues 65–170 is important for DNA binding and probably contributes to core-type DNA binding specificity.

In contrast to the gel-mobility shift assay, the following experiment showed that a region of the C170 domain is in close proximity to the DNA and provides a portion of the DNA

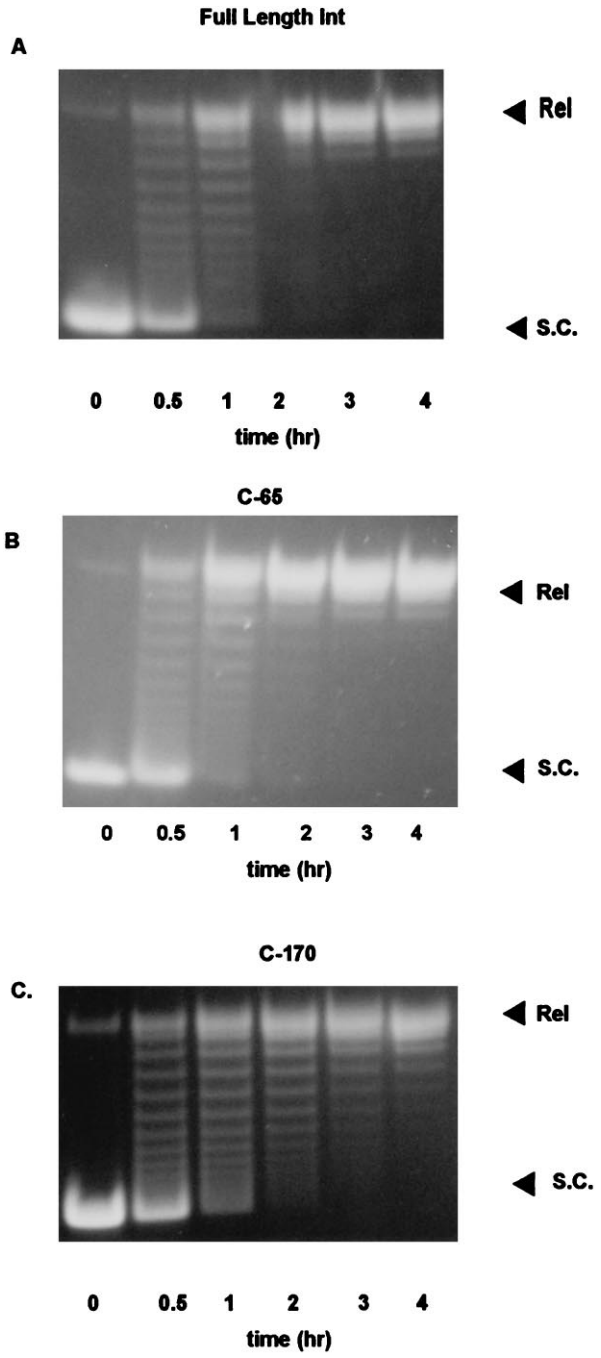

FIG. 3. Topoisomerase activities of λ Int and its carboxy-terminal fragments C65 and C170. Supercoiled pBR322 plasmid DNA (0.3 μg/μl) was incubated with 40 pmols of intact Int (*A*) or C65 (*B*) or 200 pmols of C-170 at 25°C in buffer A (20 μl of 50 mM MOPS, pH 7.4/50 mM Nacl/0.1% BSA/5 mM EDTA/1 mM DTT/10% DMSO. At the indicated times, the reactions were quenched with 0.2% SDS and electrophoresed in 1.2% agarose containing 0.5 μg/ml ethidium bromide. The electrophoresis buffer also contained ethidium bromide at a concentration of 0.5 μg/ml. The gels were destained and visualized under UV light. The positions of supercoiled (s.c.) and relaxed (Rel) plasmid DNA are indicated.

binding surface(s). A 13-bp core-type DNA fragment containing the overlap region and one core-type Int binding site, labeled at its 5′ terminus with ^{32}P, was UV-crosslinked to C65F protein. Previous experiments had shown that, under these conditions, ≈5% of the protein is crosslinked to the DNA. The mixture was then digested with Arg-C, and aliquots were analyzed by SDS/PAGE. Coomassie blue was used to visualize

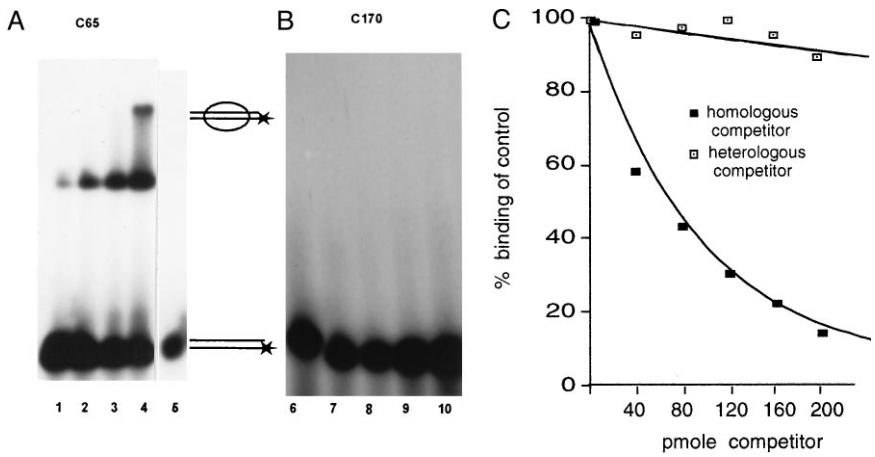

FIG. 4. Gel-mobility shift assays of C65F and C170F; 6, 12, 25, and 50 pmols of C65F (*A*, lanes 1–4) or C170F (*B*, lanes 6–9) were mixed with 25 pmols of ^{32}P-labeled synthetic 30-bp core-type DNA and incubated at room temperature for 20 min in 15 μl of buffer B (10 mM Tris, pH 7.5/50 mM NaCl/1 mM EDTA/1 mM DTT) plus 10% Ficoll. The reaction was analyzed by electrophoresis in 8% polyacrylamide (acrylamide:bis ratio 30:1). Lanes 5 and 10 show the DNA alone. (*C*) the same kind of analysis carried out under the same conditions, but in this case C65F (20 pmols) was mixed with 40 pmols of core-type DNA in the absence and presence of 40, 80, 120, 160, and 200 pmols of homologous (filled squares) or heterologous competitor (empty squares). The reaction was analyzed as above and quantitated by scanning the gels on a PhosphorImager (Molecular Dynamics).

the protein bands and indicated that most of the C65 had been cleaved to C170, as expected (Fig. 5, *Left*). Autoradiography was used to visualize the DNA and protein–DNA complexes (Fig. 5, *Right*). A significant fraction of the crosslinked DNA was found associated with the protease product, C170, indicating that residues within the C170 domain are in close contact with the DNA. They also indicate that the proteolytically sensitive site responsible for generating C170 remains exposed and accessible to protease even when the large carboxy-terminal C65 fragment is bound to *att* site DNA.

Cleavage of Half-*att* Suicide Substrates. Suicide *att* site substrates have proven to be extremely useful in studying DNA cleavage and strand transfer by Int and other site-specific recombinases. The half-*att* site version of these substrates consists of a single core-type Int binding site plus three bases of the top and seven bases of the bottom strands of the overlap region (18, 19). Upon nicking by Int, a three-base oligomer from the top strand overlap region is released and lost by diffusion, leaving Int covalently attached to the DNA (20). In the presence of a suitable recombination partner, the Int-*att* covalent complex can participate in a unitary strand transfer reaction: The 5′OH of an invading strand attacks the phosphotyrosine linkage, displacing Tyr-342 and generating a new phosphodiester linkage. In the absence of an invading 5′OH nucleophile, the phosphotyrosine bond is stable, and the covalent Int–DNA complex can be separated from free *att* site DNA by electrophoresis in an SDS polyacrylamide gel. By this assay, the C65 protein is indistinguishable from full length Int in its ability to form covalent DNA complexes (Fig. 6). The C170 fragment, at comparable protein concentration, has a very low cleavage activity (not shown), as expected from its greatly impaired DNA binding. The DNA binding defect can be overcome by increasing the protein concentration (150- to 200-fold molar excess of protein). The cleavage activity of C170 can be stimulated an additional 3- to 4-fold by 10% DMSO (as was observed for topoisomerase activity). Thus 40% of the *att* site DNA can be converted into covalent complex (Fig. 6). DMSO has no effect on the cleavage activity of the C65 fragment or the full length λ Int.

The Active Site Nucleophile Is Protease-Sensitive. In addition to the 20- and 11-kDa products generated by protease

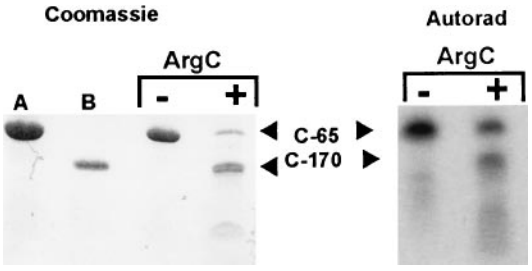

FIG. 5. UV-mediated crosslinks between core-type DNA oligomer and the carboxy-terminal domains of Int. C65 was mixed with a 5′-^{32}P-labeled, 13-bp, core-type DNA fragment (300 pmols each) in 50 μl of buffer B (see Fig. 4) plus 10% glycerol. The mixture was incubated 10 min at room temperature and 20 min on ice and then layered on a Petri dish on ice where it was irradiated with UV light (254 nm) for 20 min. An aliquot of the UV-crosslinked complex was subjected to limited proteolysis with 100 ng of endoproteinase Arg-C (rat submaxillary gland, HPLC-purified), at 25°C for 1 h and analyzed by 15% SDS/PAGE (13) followed by autoradiography. (*Left*) Coomassie blue staining of the gel before and after proteolysis. Lanes A and B contain purified recombinant C65 and C170 proteins, respectively, as markers. (*Right*) The autoradiogram of the UV-crosslinked products before and after limited proteolysis with Arg-C.

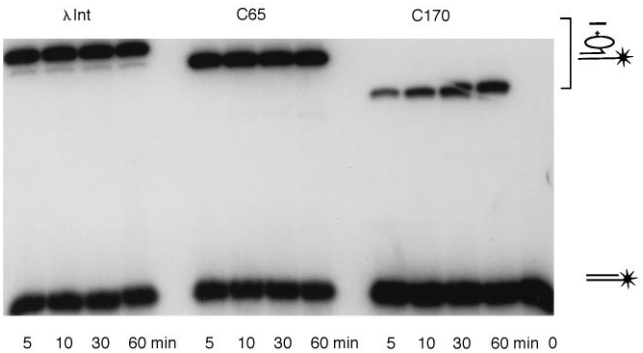

FIG. 6. Covalent complex formation between suicide *att* site substrate and Int, C65, or C170. One picomol of suicide half *att* site substrate labeled at the 5′ end of the bottom strand (as illustrated in the right margin) was incubated at 25°C with 20 pmols of intact Int or C65 (*Left* and *Center*) or 200 pmols of C170 (*Right*) in buffer A. At the indicated times, aliquots were quenched with 0.2% SDS, electrophoresed on an 15% SDS/polyacrylamide gel and visualized by autoradiography and quantitated with a PhosphorImager. The cartoon in the right margin indicates the position of the retarded covalent complexes and illustrates the loss of the three-base oligo from the 3′ end of the top strand, thus trapping the covalent complex.

cleavage just upstream of Ala-170, as described above, there was also a 19-kDa fragment. When this was subjected to N-terminal sequencing, it was found to start at Ala-170, like the 20-kDa fragment. It is smaller than C170 by ≈1.5 kDa, so this peptide was presumably generated by a second cleavage at a protease-sensitive site ≈15 residues from the carboxyl terminus. Although the expected 1.5-kDa carboxy-terminal peptide was not visualized in the gel system used in Fig. 2, the second protease-sensitive site was confirmed by the finding that purified C170 could be quantitatively converted to the 19-kDa fragment by digestion with Arg-C. To identify the protease cleavage site, an Arg-C digest of C170 was subjected to N-terminal sequencing. Two amino-terminal sequences were derived. One started with the previously identified sequence found at the amino termini of the 19-kDa and C170 peptides. The other sequence commenced with Gln-341, consistent with the protease cleavage predicted to generate the 19-kDa fragment from C170. Fragments approximately the same size as the 19-kDa fragment also were obtained with kallikrein, trypsin, and chymotrypsin digestions of C170, indicating a region extremely sensitive to protease cleavages. The 1.5-kDa fragment generated by Arg-C cleavage between Ser-340 and Gln-341 is probably the result of a nonspecific cleavage by Arg-C. Although Arg-C is predominantly specific for peptide bond cleavage at the carboxyl side of arginines, nonspecific cleavages involving Lys–Lys and Lys–Arg bonds also have been reported and used selectively (21, 22), and not all arginines are cleaved (23). Nonspecific cleavages by specific proteases (such as endoproteinase Glu-C and trypsin) are fairly common and documented (24). We do not know whether the Ser–Gln cleavage observed here is by Arg-C or by a low level contaminant in the Arg-C preparation. However, it is clear that this region, in the immediate vicinity of the nucleophilic Tyr-342, is especially susceptible to protease attack because it was uniquely sensitive to all four proteases tested.

DISCUSSION

Using the classical method of limited proteolysis, we dissected the λ Int protein and identified what we believe is the minimal catalytically competent domain. This 188-amino acid domain, extending from Ala-170 to the carboxyl terminus, has been cloned and purified to homogeneity. The C170 domain does not show efficient or specific binding to *att* site DNA in a gel retardation assay. However, it does have topoisomerase activity, and it does cleave an *att* site suicide substrate to form the covalent phosphotyrosine recombination intermediate. Consistent with its poor DNA binding, C170 requires higher concentrations of protein than C65 or intact Int to see comparable levels of "catalytic activity" (we use the term "catalytic activity" to refer to the formation of covalent Int–DNA complex, even when it does not involve turnover). The fact that 10% DMSO stimulates C170 catalytic activity 3- to 4-fold and has no affect on the activity of Int or C65 is probably due to the stimulation of C170 binding to DNA. DMSO is well known for its ability to reduce the stringency of protein–DNA interactions as is observed in the formation of transposition complexes by Mu-A (25) and the enhancement of *Eco*RI "star" activity (26). However, the role of water activity in determining the specificity of protein–DNA interactions is still the subject of much debate (27).

Based on the results reported here and previously (9), our current view is that λ Int consists of three structurally and functionally distinct domains. The amino-terminal domain, from Met-1 to Leu-64, specifically recognizes distal arm-type DNA sequences. The 11-kDa domain, from Thr-65 to (presumably) Arg-169, is obviously important for specific binding to core-type DNA sequences at the sites of strand exchange, and its capacity for autonomous function is currently under investigation. It has also been suggested that this region may be involved in protein–protein interactions (28). The C170 carboxy-terminal domain includes all of the catalytic machinery necessary for DNA cleavage and ligation.

Another member of the Int family, Flp recombinase (FLP) protein from yeast, also has been analyzed by limited proteolysis, and peptides consisting of the carboxyl two-thirds of FLP have been shown to bind to the FRT target site (29, 30). An amino-terminal peptide, residues 2–123, has been shown to possess nonspecific DNA binding activity and to enhance the specific binding of the carboxy-terminal peptides to the FRT sites (31). Additionally, the cleavage and ligation activities could be demonstrated by reconstituting FLP with a combination of peptides 2–123 and 124–423 (31).

The catalytically active C170 λ Int domain contains all of the residues that have been identified as being conserved in the Int family of recombinases (4–6, 32, 33). This includes the very highly conserved triad of Arg-212, His-308, and Arg-311 that has been suggested to activate the scissile phosphate for DNA cleavage (34, 35), the active site nucleophile Tyr-342, (20), and Glu-174, which can be mutated to give a hyperrecombination phenotype (36). The size of the C170 domain is approximately equal to that of smallest Int family members, such as FimB and FimE of *E. coli*, 227 and 209 amino acids, respectively (37). We also note that, when Int and FLP are aligned, two of the protease-sensitive sites in FLP (30, 31) flank the λ 170 position with intervals of 5 and 15 amino acids, respectively [using the alignment in Kwon *et al.* (6)].

In addition to the protease-sensitive site that generated the C170 domain, we also have noted that Tyr-342, the active site nucleophile, is in a uniquely protease-sensitive region, suggestive of a flexible loop for this portion of the protein. This suggestion is borne out by the x-ray crystal structure of C170 that has recently been solved at a 1.9-Å resolution by Kwon *et al.* (6). In this structure, Tyr-342 is seen to be on a 17-amino acid loop extending from Lys-334 to Trp-350. It has been suggested that this flexible loop with the catalytic Tyr is a general feature of the Int family of recombinases and that it may explain the duality of cleavage mechanisms that is observed within the Int family (for reviews of *cis* vs. *trans* cleavage mechanisms, see refs. 1, 2, 34, 38). A similar sensitivity to protease cleavage is seen in the immediate vicinity of the catalytic Tyr of FLP. We note that no evidence was found for protease-sensitive sites in any other regions of λ Int known to be critical for catalytic function, e.g., Arg-212 or His-308/Arg-311.

Comparison of the C170 crystal structure with a tree-based algorithm alignment of the Int family recombinases showed that (*i*) the invariant and catalyticaly critical residues of the Int family are all located on the proposed DNA interaction surface of λ C170; (*ii*) few of the exposed residues away from the proposed active site are conserved in the Int family; and (*iii*) many of the buried residues of C170 are conserved (6). Based on the results presented here and the observations from the C170 crystal structure, we suggest that a "C170 domain" is likely to exist for all, or a majority, of the Int family recombinases. The isolated catalytic domain is well suited for many types of mechanistic studies, and it is not unreasonable to anticipate that different recombinases from the 70-member Int family will be the source of different aspects of the chemistry of this site-specific recombination pathway.

We are greatly indebted to: Istvan Papp and Matxalen Llosa for some of the initial constructions in cloning C65Int; David Woo for analyses of C65 binding to core-type DNA fragments; Tina Oliveira for technical assistance; Joan Boyles for preparation of the manuscript; and members of the Landy laboratory for advice and helpful comments throughout this project. This work was supported by National Institutes of Health Grants AI13544 and GM33928.

1. Landy, A. (1993) *Curr. Opin. Genet. Dev.* **3,** 699–707.

2. Sadowski, P. D. (1993) *FASEB J.* **7,** 760–767.
3. Stark, W. M., Boocock, M. R. & Sherratt, D. J. (1992) *Trends Genet.* **8,** 432–439.
4. Leong, J. M., Nunes-Düby, S., Oser, A., Youderian, P., Susskind, M. M. & Landy, A. (1984) *Cold. Spring Harbor Symp. Quant. Biol.* **49,** 707–714.
5. Abremski, K. E. & Hoess, R. H. (1992) *Protein Eng.* **5,** 87–91.
6. Kwon, H. J., Tirumalai, R., Landy, A. & Ellenberger, T. (1997) *Science* **276,** 126–131.
7. Campbell, A. M. (1962) in *Episomes*, ed. Caspari, E. W. (Academic, New York), pp. 101–145.
8. Kikuchi, Y. & Nash, H. A. (1978) *J. Biol. Chem.* **253,** 7149–7157.
9. Moitoso de Vargas, L., Pargellis, C. A., Hasan, N. M., Bushman, E. W. & Landy, A. (1988) *Cell* **54,** 923–929.
10. Henikoff, S. (1987) *Methods Enzymol.* **155,** 156–165.
11. Tabor, S. & Richardson, C. C. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 1074–1078.
12. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) in *Molecular Cloning* (Cold Spring Harbor Lab. Press, Plainview, NY).
13. Laemmli, U. K. (1970) *Nature (London)* **227,** 680–685.
14. Schulz, G. & Schrimer, R. (1979) in *Principles of Protein Structure* (Springer, New York).
15. Janin, J. & Chothia, C. (1985) *Methods Enzymol.* **115,** 420–430.
16. Kikuchi, Y. & Nash, H. A. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 3760–3764.
17. Hedrick, J. L. & Smith, A. J. (1968) *Arch. Biochem. Biophys.* **126,** 155–164.
18. Nunes-Düby, S. E., Matsumoto, L. & Landy, A. (1989) *Cell* **59,** 197–206.
19. Nunes-Düby, S., Azaro, M. & Landy, A. (1995) *Curr. Biol.* **5,** 139–148.
20. Pargellis, C. A., Nunes-Düby, S. E., Moitoso de Vargas, L. & Landy, A. (1988) *J. Biol. Chem.* **263,** 7678–7685.
21. Bousfield, G. R. & Ward, D. N. (1988) *J. Biol. Chem.* **263,** 12602–12607.
22. Chartier, F., Laine, B., Belaiche, D. & Sautiere, P. (1989) *J. Biol. Chem.* **264,** 17006–17015.
23. Moonen, P., Akeroyd, R., Westerman, J., Puijk, W. C., Smits, P. & Wirtz, K. W. (1980) *Eur. J. Biochem.* **106,** 279–290.
24. Keil, B. (1982) in *Methods in Protein Sequence Analysis* (Humana, Clifton, NJ), pp. 291–304.
25. Baker, T. A. & Mizuuchi, K. (1992) *Genes Dev.* **6,** 2221–2232.
26. Robinson, C. R. & Sligar, S. G. (1993) *J. Mol. Biol.* **234,** 302–306.
27. Sidorova, N. Y. & Rau, D. C. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 12272–12277.
28. Han, Y. W., Gumport, R. I. & Gardner, J. F. (1994) *J. Mol. Biol.* **235,** 908–925.
29. Pan, G., Clary, D. & Sadowski, P. D. (1991) *J. Biol. Chem.* **266,** 11347–11354.
30. Chen, J.-W., Evans, B. R., Yang, S.-H., Teplow, D. B. & Jayaram, M. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 5944–5948.
31. Pan, G. & Sadowski, P. D. (1993) *J. Biol. Chem.* **268,** 22546–22551.
32. Argos, W., Landy, A., Abremski, K., Egan, J. B., Haggard-Ljungquist, E., Hoess, R. H., Kahn, M. L., Kalionis, W., Narayana, S. V. L., Pierson, L. S. I., Sternberg, N. & Leong, J. M. (1986) *EMBO J.* **5,** 433–440.
33. Blakely, G. W. & Sherratt, D. J. (1996) *Mol. Microbiol.* **20,** 234–237.
34. Chen, J., Lee, J. & Jayaram, M. (1992) *Cell* **69,** 647–658.
35. Pan, G., Luetke, K. & Sadowski, P. D. (1993) *Mol. Cell. Biol.* **13,** 3167–3175.
36. Lange-Gustafson, B. J. & Nash, H. A. (1984) *J. Biol. Chem.* **259,** 12724–12732.
37. Klemm, P. (1986) *EMBO J.* **5,** 1389–1393.
38. Stark, W. M. & Boocock, M. R. (1995) *Trends Genet.* **11,** 121–123.