Cleavage Properties of an Archaeal Site-specific Recombinase, the SSV1 Integrase*

Received for publication, January 23, 2002, and in revised form, February 21, 2002
Published, JBC Papers in Press, March 1, 2002, DOI 10.1074/jbc.M200707200

Marie-Claude Serre‡§, Claire Letzelter¶, Jean-Renaud Garel†**, and Michel Duguet‡

From the ‡Laboratoire d’Enzymologie et Biochimie Structurales, CNRS Bat. 34, avenue de la Terrasse, 91198 Gif-sur-Yvette Cedex, France and the §Laboratoire d’Enzymologie des Acides Nucléiques, Institut de Génétique et Microbiologie bât 400, Université Paris Sud, 91405 Orsay Cedex, France

SSV1 is a virus infecting the extremely thermophilic archaeon Sulfolobus shibatae. The viral-encoded integrase is responsible for site-specific integration of SSV1 into its host genome. The recombinant enzyme was expressed in Escherichia coli, purified to homogeneity, and its biochemical properties investigated in vitro. We show that the SSV1 integrase belongs to the tyrosine recombinases family and that Tyr114 is involved in the formation of a 3’-phosphotyrosine intermediate. The integrase cleaves both strands of a synthetic substrate in a temperature-dependent reaction, the cleavage efficiency increasing with temperature. A discontinuity was observed in the Arrhenius plot above 50 °C, suggesting that a conformational transition may occur in the integrase at this temperature. Analysis of cleavage time course suggested that noncovalent binding of the integrase to its substrate is rate-limiting in the cleavage reaction. The cleavage positions were localized on each side of the anticodon loop of the tRNA gene where SSV1 integration takes place. Finally, the SSV1 integrase is able to cut substrates harboring mismatches in the binding site. For the cleavage step, the chemical nature of the base in position −1 of cleavage seems to be more important than its pairing to the opposite strand.

Site-specific recombination catalyzed by tyrosine recombinases plays a number of critical roles in prokaryote and eukaryote kingdoms. Well documented examples in lower eukaryotes and bacteria include generation of genetic variability, plasmid copy control and/or stable inheritance, resolution of bacterial chromosome dimers or viral DNA integration in host chromosomes (for reviews, see Refs. 1–3). Members of the tyrosine recombinases family catalyze site-specific recombination between two DNA sites by using a topoisomerase IB-like mechanism to cut and religate DNA strands (4, 5). Unlike topoisomerases, tyrosine recombinases perform the ligation step after strand exchange between the two DNA partners. Site-specific recombination requires the assembly of a synaptic complex containing, at least, four enzyme protomers and the two DNA sites. The recombination reaction occurs by cutting and exchanging the two pairs of DNA strands in two temporally distinct steps. In the first step, cleavage occurs on the top strands of each DNA site. For each site, a 3’-phosphotyrosine DNA-protein covalent complex is formed and a free 5’-OH DNA end generated. The leaving strands then attack the phosphotyrosine link of the recombination partner, thus releasing the recombinase subunits. After this first round of strand cleavage-strand exchange, a Holliday junction is formed. The bottom strands are then cut and religated, thus resolving the Holliday junction and completing the recombination reaction.

Site-specific recombination in archaea is not as well known. So far, the only studied system is SSV1, a virus of the extremely thermophilic archaeon Sulfolobus shibatae. In the cell, the 15.5-kb genome of SSV1 is present both as a circular DNA and as a provirus stably integrated into an arginine tRNA gene (6, 7). Site-specific integration of SSV1 into its host chromosome is catalyzed by the virus-encoded integrase (IntSSV). This enzyme catalyzes recombination between viral and chromosomal attachment sites, attP and attB (the latter previously denoted attA), to generate a left (attL) and right (attR) prophage att sites (7, 8). In all four att sites, a 44-bp invariant sequence (previously referred as to the core sequence (7)) is found. This 44-bp segment comprises the 3’-half of the arginine tRNA gene and flanks the provirus as direct repeats (7). In vitro, IntSSV was shown to recombine DNA substrates both in the integrative and excisive pathways (8, 9).

Based on sequence alignments, IntSSV was proposed to be a member of the tyrosine recombinases family (10). One striking aspect of this family is the lack of global homology among its more than 130 members. Nevertheless, a conserved signature is found in the C-terminal part of all the proteins. All members of the family harbor two short regions of similarity, box I and box II, sharing an invariant RHRY amino acids tetrad (11–14). In different systems, mutations introduced at any of these positions produced proteins deficient in recombination, as expected for active site residues. The conserved tyrosine of the tetrad is the nucleophilic group which attacks the scissile phosphate of each DNA site, while the conserved RHR are involved in the correct positioning of the phosphate group and behave as a charge relay system within the catalytic pocket. For IntSSV, Tyr114 was proposed to be the catalytic residue (10). However, there was no biochemical evidence that IntSSV would catalyze recombination in the same way as other members of the family. Furthermore, the observations that IntSSV is the most distantly related member of the tyrosine recombinases family (13) and that it harbors substitutions at several conserved positions (14), raised the possibility that the recombination mechanism might be different in archaea.

To gain more insights about the IntSSV recombination mechanism, we cloned the wild type IntSSV gene, as well as a Tyr114 to Phe mutant (IntSSV-Y314F), in an E. coli expression vector. We developed a purification procedure that yields homoge-
neously purified untagged recombinant enzymes. The present work focuses onto the first stage of catalysis which leads to the cleavage of the DNA substrate. We show that, in vitro, the SSV1 integrase is able to cut a double-stranded synthetic substrate harboring a specific target sequence. The efficiency of the cleavage reaction is strictly dependent on temperature. Analysis of the reaction kinetics suggests a rapid equilibrium between cleavage and recombination within the integrase-substrate complex, a feature that can be expected for a topoisomerase IB-like mechanism. Moreover, the absence of substrate cleavage when using the IntSSV-Y314F mutant strongly suggests that like other members of the tyrosine recombinases family, IntSSV mediates cleavage through a 3'-phosphotyrosine intermediate. Finally, we have identified the cleavage points on the double-stranded substrate and found that they correspond to the borders of the anticond loop of a tRNA arginine gene. This finding is reminiscent of the integration of some bacterial phages and plasmids into tRNA genes (15). This is the first case described in archaea, suggesting that targeting of tRNA genes is an ancient process that was conserved during evolution of bacteriophages.

MATERIALS AND METHODS

Bacterial Strains and Media—The Escherichia coli K12 strains used were as follows. C600 recA was used for cloning experiments and DNA amplification, except for the T-tailing vector and derivatives which were transformed in strain Mos Blue (Amersham Pharmacia Biotech). MC1061 (16) was transformed by plasmid pSG3 (described below) to give the integrase producing strain, MS91L. When necessary, Luria-Bertani medium was supplemented with ampicillin (100 μg/ml) or tetracyclin (12.5 μg/ml).

DNA Techniques—Large scale or minipreparations of plasmid DNA were performed as described previously (17). Restriction and modification were as described by the manufacturer. Purification of DNA fragments (GeneClean, Bio 101), cloning of PCR fragments in a T-tailing vector (pMOsBlue T-vector kit, Amersham Bioscience), and DNA sequencing (Thermo Sequenase cycle sequencing kit, Amersham Bioscience) were done according to the supplier’s instructions. Oligonucleotides were purchased from Genset and MWG.

Cloning of the Native IntSSV Gene under the Arabinose Promoter—Plasmid pBAD18, a pBR322 derivative carrying the promoter of E. coli and its regulator gene araC (18), was modified to be used as the expression vector. The unique Ndel site of pBAD18 was removed by fill in (pIG1). A XbaI-HindIII fragment containing the Shiga-Dalgaro sequence and initiating ATG (overlapping a Ndel site) from gene 010 of bacteriophage T7 was then subcloned into the corresponding sites of pSG1 (pSG2). The IntSSV coding sequence was amplified by PCR using pBAD18 as template and the following primers 5’-CGCTTCGAACCTTTTACGCTTT-3’ (MUT3). The underlined positions correspond to the Ndel and HindIII sites created. In bold are shown the initiating codon (Ndel) and stop codon (MUT3). Italics indicate the mutations introduced. The resulting 1,025-bp fragment was cloned into the pMOsBlue vector, and its sequence verified. The Ndel-HindIII 1,010-bp fragment was then subcloned into the corresponding sites of pSG2, to give the expression plasmid pSG3.

Site-directed Mutagenesis of the Gene Encoding IntSSV—The 1,049-bp XbaI-HindIII fragment containing the IntSSV coding sequence was subcloned into the corresponding restriction sites of M13mp18 (19). Site-directed mutagenesis was performed on the uracil-containing template of the resulting plasmid as described previously (20) but with T7 DNA polymerase for the elongation step. The mutation (underlined position) Y314F was introduced by using the following oligonucleotide: 5’-GAGG-GATACGAAATGTTGCG-3’ (GAGG-GATACGAAATGTTGCG-3’-MUT5). The nucleotide sequence of the mutated fragment was then verified by sequencing prior to recloning in pSG3.

Expression and Purification of IntSSV—An overnight culture of strain MS91L in Luria-Bertani, ampicillin, 0.2% glucose, was diluted to A600 = 0.07 in Luria-Bertani, ampicillin. The culture was grown at 37 °C to A600 = 0.6 and expression of IntSSV was induced by the addition of arabinose (1% final concentration). 2 h after induction (A600 = 1.1), cells were harvested by centrifugation. The cell pellet was washed with TN buffer (10 mM Tris, 100 mM NaCl) and centrifuged again. The cell pellet was resuspended in 5 ml of A500 buffer (50 mM Tris, pH 7.5, 100 mM NaCl, 1 mM EDTA, 5 mM β-mercaptoethanol) per g of wet cells. Lysozyme was added (1.7 mg/g of cells), and incubation performed for 1 h at 4 °C. 3 cycles of freezing/thawing were then performed, and the resulting crude extract was ultracentrifuged 45 min at 4 °C and 200,000 × g. The supernatant was collected (fraction F1) and submitted to fast protein liquid chromatography (fraction F2) on a MonoQ 5/50 column equilibrated in 1.1 buffer. The protein was eluted with a linear NaCl gradient (150 mM NaCl, 1 M NaCl). The protein peak was pooled and analyzed by SDS-PAGE. The protein peak was dialyzed overnight at room temperature against 500 ml of A500 buffer. The dialyzed sample was collected and centrifuged at room temperature 10 min at 10,000 × g. The supernatant (fraction FII) was collected and heated for 10 min at 65 °C. After centrifugation at 10,000 × g, the sample (fraction FIV) was further purified by size exclusion chromatography performed at room temperature, onto a Sephacryl S300 column (Amersham Bioscience) equilibrated in A500 buffer. The peak fractions containing the integrase were pooled and concentrated in a Vivaspin concentrator (MWCO 30,000), then dialyzed overnight against A500 buffer (same as A500, but 500 mM NaCl and 20% glycerol). The protein (fraction FV) was either stored at 4 °C for up to 1 month or at −70 °C for up to 6 months. The yield of purification is of 450 μg of pure IntSSV per liter of bacterial culture.

Fluorescence—Protein fluorescence was followed with a Bio-Tek SFM 25 fluorescence. Fluorescence measurements were carried out in a buffer composed of 50 mM Tris, pH 7.5, 125 mM NaCl, 1 mM EDTA, containing 1 mM dithiothreitol. Protein concentration was 0.5 or 2.5 μM and the protein was incubated with excitation at 4°C and emission spectra, respectively. Emission spectra of the protein excited at 295 nm were recorded at different temperatures.

Cleavage Assays—The cleavage reactions were carried out in 20 μl of reaction mixtures consisting of 30 mM Tris, pH 7.5, 50 μM bovine serum albumin, 125 mM NaCl, and the indicated amounts of 5°-end labeled DNA substrate. The reaction mixtures were incubated at 65 °C (unless noted otherwise) and at the times indicated quenched in Laemmli loading buffer (final concentrations: 40 mM Tris, pH 6.8, 3% SDS, 8% glycerol, 250 mM β-mercaptoethanol, 0.005% bromphenol blue) and heated for 5 min at 98 °C. The reaction products were then analyzed by electrophoresis through 12% SDS-polyacrylamide gel. The bands were autoradiographed, and bands corresponding to the DNA complex and substrate DNA were quantitated using the Scan Analysis 2.11 software. The results presented are the mean of at least three independent experiments.

Alternatively, a nonadenaturing loading buffer (final concentrations: 2 mM Tris, pH 7.5, 0.2 mM EDTA, 4% glycerol, 20 μg/ml bovine serum albumin, 200 μg/ml xylene cyanol) was added to the cleavage reactions, and the samples were analyzed on 8% polyacrylamide gels (30/0.5 acrylamide:bisacrylamide) in TGE buffer (50 mM Tris, 8 mM glycin, 0.1 mM EDTA). Samples were electrophoresed at 4°C for 3 h at 7 V/cm.

Determination of the cleavage position was performed by incubating 2 μg IntSSV with 50 mM 3°-end labeled double-stranded substrate for 4 h at 65 °C in the same buffer conditions. The reaction was stopped by heating for 5 min at 95°C in 5% deionized formamide, 10 mM EDTA, 0.3% bromophenol blue, 0.3% xylene cianol blue, and the sample heated for 5 min at 98 °C. The reactions were analyzed on a 18% polyacrylamide gel (19:1) containing 8 M urea, in TBE buffer (90 mM Tris, 90 mM boric acid, 1 mM EDTA). Electrophoresis was performed at 52 V/cm. The cleavage positions were determined by comparing electrophoretic mobilities of the samples to that of specific and nonspecific ladders (see “Results”).
RESULTS AND DISCUSSION

Expression and Purification of the Recombinant Int<sup>SSV</sup>—

Cloning of the Int<sup>SSV</sup> gene (SSV1 open reading frame D-335) was first described by Muskhelishvili et al. (8). However, the strategy of cloning into pGEX-2T expression vector led to the addition of 2 residues at the N terminus and 7 residues at the C terminus of the protein. To investigate the biochemical properties as well as to start up crystallographic studies, we wanted to express the untagged integrase protein. The cloning strategy of the Int<sup>SSV</sup> gene into a pBAD vector is described under Materials and Methods. When induced with arabinose, strain MSG91 produced a recombinant Int<sup>SSV</sup> with a molecular mass on SDS-PAGE close to the predicted value of 38.9 kDa. Under the expression conditions described under “Materials and Methods,” Int<sup>SSV</sup> accounts for 10–15% of total protein mass.

The purification scheme of the recombinant Int<sup>SSV</sup> takes advantage of two intrinsic properties of the enzyme, its affinity for nucleic acids and its thermostability, so that the enzyme could be easily purified to near homogeneity. At low ionic strength in the presence of nucleic acids, Int<sup>SSV</sup> remains in the soluble fraction (fraction I) after centrifugation at 200,000 × g (Fig. 1, lane 1). When adding streptomycin sulfate (2% final) to fraction I, Int<sup>SSV</sup> co-precipitates with nucleic acids. About 80% of the proteins are eliminated during this fractionation step, while most of Int<sup>SSV</sup> remains in the nucleic acids pellet. The pellet is resolubilized in a high ionic strength buffer (A 1.0) to give fraction II (Fig. 1, lane 2). Remaining RNA contaminants are eliminated by RNase A treatment of fraction II. After dialysis, the sample (Fig. 1, lane 3) is submitted to thermal denaturation for 10 min at 65 °C, and the insoluble material removed by centrifugation. This step eliminates 65% of the remaining contaminant proteins. The collected soluble fraction (Fig. 1, lane 4) is injected onto a Sephacryl S300 size exclusion column equilibrated in A<sub>1.0</sub> buffer. Int<sup>SSV</sup> elutes in a single peak, with an apparent molecular mass of 42.7 kDa indicating that, under these conditions, the protein behaves as a monomeric species. The peak fractions are then pooled and concentrated by ultrafiltration (Fig. 1, lane 5) to a final concentration of 0.5 to 1 mg/ml. The purified integrase can be stored at 4 °C for up to 1 month or at −70 °C for up to 6 months without significant loss of cleavage activity. The recombinant Int<sup>SSV</sup> purified from E. coli retains its thermostability and affinity for nucleic acids, suggesting that, even though expression is performed at a low temperature, the recombinant enzyme is correctly folded.

Int<sup>SSV</sup> Is Able to Cleave a Minimal Substrate in Vitro, Forming a Covalent Link via Tyr<sup>314</sup>—

The reaction catalyzed by tyrosine recombinases involves a transient protein-DNA covalent intermediate between a DNA 3’-phosphoryl group and the conserved tyrosine residue. This mechanism is supposed to be conserved within the family. Therefore, the prediction on Int<sup>SSV</sup> activity was that the enzyme would be able to cleave a specific DNA substrate and to form a 3’-phosphoprotein covalent intermediate.

To monitor the cleavage reaction, we designed a 19-bp double-stranded synthetic substrate (XTB, Table I). XTB contains the 18 bp of the attP site protected by Int<sup>SSV</sup> (9). Int<sup>SSV</sup> (1 or 2 μM) was incubated 4 h at 65 °C with the XTB substrate (12.5 nM), in conditions described under “Materials and Methods.” After heat denaturation, the reaction products were analyzed by SDS-PAGE (Fig. 2A). The appearance of a low mobility product (lanes 2 and 4) indicates that a covalent complex is formed between the integrase and the substrate. The same result is obtained when the bottom strand of the substrate is 5’-end labeled (Table I). Transfer of the labeling to the enzyme reveals that it is covalently bound to the 3’-DNA end, confirming that the polarity of strand cleavage by Int<sup>SSV</sup> is the same as that of other members of the tyrosine recombinases family.

The specificity of the cleavage reaction was further assessed by adding increasing amounts of nonspecific DNA (Table I). Cleavage on the top and bottom strands was not significantly reduced in the presence of 0.5 μg of poly(dI-dC)<sub>9</sub> used as nonspecific competitor DNA. These results indicate that the reaction is specific to the XTB-encoded sequence since a large excess of nonspecific DNA (ratio site:competitor in bp of 1 to 150) has very little effect on the cleavage efficiency. The XTB substrate thus contains enough information to allow specific cleavage by Int<sup>SSV</sup> on the top and bottom strands.

Alignment of Int<sup>SSV</sup> amino acid sequence with the other members of the family predicted that Tyr<sup>314</sup> is the active site tyrosine (10, 13, 14). To demonstrate that Tyr<sup>314</sup> is indeed involved in the covalent complex formation, the activity of the mutant Int<sup>SSV</sup>_Y314F has been investigated. The purification procedure applied to the mutant protein gave identical results to that obtained with the wild type enzyme (data not shown). As presented in Fig. 2A (lanes 3 and 5), the Int<sup>SSV</sup>_Y314F mutant is totally inactive in the cleavage reaction, indicating that the hydroxyl group of Tyr<sup>314</sup> is essential for the formation of the covalent complex. Thus, like bacterial and eukaryotic tyrosine recombinases, the archaeal integrase mediates cleavage via a 3’-phosphotyrosine intermediate.

Determination of the Cleavage Position within the Synthetic Substrate—in Vivo, SSV1 specifically integrates into a tRNA<sup>Asp</sup> gene either in S. shibatae or Sulfolobus solfataricus (7, 23). In bacteria, several phages and plasmids encoding an integrase of the tyrosine recombinases family are also integrating into tRNA genes. Examples include phages P4 from E. coli (24), HP1 from Haemophilus influenzae (25), MV4 from Lactobacillus delbrueckii subsp. bulgaricus (26), L5 from M. smegmatis (27), or plasmid pSAM2 from Streptomyces ambofaciens (28).

It was suggested that tRNA genes may have been the recognition sequences for an ancestral site-specific recombination enzyme (7). It was also proposed that “the frequent use of tRNA genes as insertion sites relates to the suitability of anticondor stem sequences as elements for core sites” (15). Based on these assumptions, we predicted that the site specificity of Int<sup>SSV</sup> would lead to strand cleavages on each side of the anticondor loop (Fig. 2B).

To identify the Int<sup>SSV</sup> cleavage sites, the XTB substrate was 3’-end labeled on its top or bottom strands. After incubation with the enzyme, the reactions were analyzed on a denaturing
DNA Cleavage Activity of the SSV1 Integrase

The sequence in brackets is the 44-bp core sequence, nucleotides underlined correspond to the anticodon loop sequence of the tRNA<sup>Arg</sup> (7). The bold type characters indicate the H<sub>2</sub>O<sub>2</sub> protected sequence (9).

Table I: Covalent complex formation between Int<sup>SSV</sup> and different synthetic substrates

| Substrate name | Labeled strand | % of covalent complex<sup>a</sup> formed | % of covalent complex formed in presence of poly(dI-dC)<sub>2</sub><sup>b</sup> |
|----------------|---------------|------------------------------------------|--------------------------------------------------|
| XTB            | Top           | 100                                      | 96                                               |
|                | Bottom        | 96                                       | 94                                               |
| XT             | Top           | 100                                      | 96                                               |
|                | Bottom        | 66                                       | 22                                               |
| XB             | Top           | 42                                       | ND<sup>c</sup>                                    |
|                | Bottom        |                                          |                                                  |

<sup>a</sup> All substrates were 5'-end labeled. The amount of covalent complex formed was quantified by scanning of the autoradiograms (covalent = unreacted = 100%) then normalized with cleavage of the XTB top strand taken as 100% (original covalent product formed after 4 h at 65 °C, 38%).

<sup>b</sup> 0.5 µg of poly(dI-dC)<sub>2</sub> added in the reaction as nonspecific competitor.

<sup>c</sup> ND, not done.

The sequence of the XTB site is shown, with bases corresponding to the 44-bp core sequence (7) in relief characters. Bases in bold type italics correspond to the predicted B and B' core type sites (see also Fig. 3B). The diagonal arrows indicate the positions cleaved by Int<sup>SSV</sup>. The region between these two arrows defines the anticodon loop, C, determination of the cleavage positions on the XTB substrate. Int<sup>SSV</sup> (2 µM) was incubated for 4 h at 65 °C with 50 nM 3'-end labeled XTB site. Labeling was done on the top strand (XT<sup>°</sup>) at both strand (XTB<sup>°</sup>). The sequence of the XTB site is shown at the bottom, with the anticodon loop sequence underlined. Positions cleaved by Int<sup>SSV</sup> on the top (CT) and bottom (CB) strands are indicated by arrows. The sequences of the XT and XB cleavage markers (14 and 12 nucleotides, respectively) are also indicated. The extra nucleotide in 3'-end labeled xtal (bold type italics) corresponds to the labeled ddA incorporated. M, 8–32 nucleotide ladder; lane 1, XT<sup>°</sup> or no Int<sup>SSV</sup>; lane 2, XT<sup>°</sup> + Int<sup>SSV</sup>; lane 3, 14-nucleotide XT<sup>°</sup> cleavage marker; lane 4, XT<sup>°</sup> + no Int<sup>SSV</sup>; lane 5, XT<sup>°</sup> + Int<sup>SSV</sup>; lane 6, 12-nucleotide XB<sup>°</sup> cleavage marker. The 8–32-nucleotide oligo ladder was 5'-end labeled, while the two cleavage markers were 3'-end labeled.

Fig. 2. Int<sup>SSV</sup> specifically cleaves synthetic substrates. A, covalent complex formation between the XTB substrate and Int<sup>SSV</sup> or Int<sup>SSV</sup>-Y314F. 12.5 nM XTB 5'-end labeled on the top strand were incubated for 4 h at 65 °C with wild type (WT) or mutant (Y314F) Int<sup>SSV</sup>. After denaturation, samples were electrophoresed on a 12% SDS-polyacrylamide gel and visualized by autoradiography. The free substrate and unreacted 5'-end labeled substrate. This behavior correlates the

gel (Fig. 2C). Preliminary experiments revealed that the oligonucleotides used as minimal substrates had abnormal migration behavior either on native or denaturing gels. Therefore, to unambiguously localize cleavage positions, two kinds of radio-labeled size standards were used: a nontpecific size ladder ranging from 8 to 32 nucleotides (Fig. 2C, lanes M) and two specific oligonucleotides corresponding to the expected products of cleavage on the top (14 nucleotides) and bottom (12 nucleotides) strands (Fig. 2C, lanes 3 and 6). Analysis of the reactions (Fig. 2C, lanes 2 and 5) reveals that Int<sup>SSV</sup> cuts XTB at specific positions which are located on each side of the anticodon loop sequence (Fig. 2B). 3'-End labeling produces a blunt end substrate which has a cleavage efficiency lower than that of a 5'-end labeled substrate. This behavior correlates the
yields obtained when using 5'-end labeled blunt end substrates. The same cleavage positions were observed when a longer substrate (FTB, Table I) was used, although with a lower efficiency (data not shown). These results indicate that XTB has all the minimal sequences required by IntSSV to get cleavage specificity.

As described for the pSAM2 integrase (28) and for the HP1 and L5 integrases (25, 27), the specificity of insertion is directed toward the anticodon loop. This result strengthens the prediction made by Campbell (15) that when attB sites overlap a tRNA anticodon stem-loop, the anticodon loop coincide with the exchange (O) region, while the sequences of dyad symmetry composing the tRNA anticodon stem are likely to be the integrase core-binding sites from the left and right parts of attB and attP (named B, B', P, P') and are aligned and the derived consensus sequence is at the bottom. Different strands annealing generate mismatched substrates. Possible annealing of XT and XB oligonucleotides correspond to the BspEl restriction site and arrows indicate the IntSSV cleavage positions. CT and CB refer to the top and bottom strand cleavage, respectively.

The IntSSV-XTB Complex Is in Equilibrium between Covalent and Noncovalent Species—The extent of XTB cleavage was studied, revealing that the maximal activity is reached for a molar ratio of 40:1 (enzyme:substrate), whatever strand is cleaved (not shown). We then followed the time course formation of the covalent complex between IntSSV and the XTB substrate at 65°C for 24 h for a fixed XTB and different enzyme concentrations (Fig. 4A), with an excess of enzyme over its substrate. At the lowest enzyme concentration of 0.05 μM, the fraction f_{cov} of substrate covalently bound to the integrase still increased after 24 h. For 0.1 or 0.25 μM IntSSV, f_{cov} reached a constant value around 0.38 to 0.4 after 24 h. For the highest enzyme concentrations of 1 and 2 μM, f_{cov} reached the same final value of 0.4 after only 2 h, and did not increase further. This showed that at most 40% of the substrate could establish a covalent link with the integrase, although there was a large

![Diagram A](image)

**Fig. 3. IntSSV-binding sites.** A, comparison of the core sequences of attB and attP. Only the top strand sequences are represented and aligned by the cleavage sites (vertical arrows). The boxed sequences correspond to the tRNA anticodon loop. The horizontal arrows above the attB sequence indicate the 10-nucleotide long imperfect inverted repeats at the attB site. The horizontal line between the attB and attP sequences represents the region protected by IntSSV from H2O2 cleavage (9). B, alignment of the four core type sequences. Sequences from attB (B, B') and attP (P, P') are aligned and the derived consensus sequence is at the bottom. C, different strands annealing generate mismatched substrates. Possible annealing of XT and XB oligonucleotides correspond to the BspEl restriction site and arrows indicate the IntSSV cleavage positions. CT and CB refer to the top and bottom strand cleavage, respectively.

1 M. C. Serre, unpublished data.
molar excess of enzyme (as high as 160-fold). It was striking that the final value of $f_{cov}$ was (almost) the same, independently of the total enzyme concentration $E_0$ in the range from 0.1 to 2 $\mu$M (Fig. 4A). This suggested to us that the covalently bound substrate was not in direct equilibrium with free enzyme. In the different assays the total integrase $E_0$ (between 0.05 and 2 $\mu$M) was in molar excess over the substrate (12.5 nM), so that, above 0.1 $\mu$M integrase, the concentrations of free and total integrase were very close.

The simplest scheme describing this situation is,

\[
E + S \xleftrightarrow{k_1} ES \xrightarrow{k_{-1}} \text{ES}_{cov}
\]

where $ES$ and $\text{ES}_{cov}$ are, respectively, a noncovalent and a covalent complex between the integrase and its substrate. Scheme 1 assumes that the cleavage step is at equilibrium relative to the binding step, and thus that binding of the substrate is concomitant with its cleavage. The alternative model,

**Fig. 4.** Dependence of the cleavage reaction upon Int$^{SSV}$ concentration. A, time course of the cleavage reaction. Different concentrations of Int$^{SSV}$ were incubated at 65 °C with XTB (12.5 nM). The reactions were stopped at different times and the samples analyzed on a 12% SDS-polyacrylamide gel. Quantification of the covalent complexes formed was done by scanning the autoradiogram of the gel. Int$^{SSV}$ concentrations: □, 0.05 $\mu$M; ■, 0.1 $\mu$M; ○, 0.25 $\mu$M; △, 1 $\mu$M; ▲, 2 $\mu$M. B, complex formation between Int$^{SSV}$ and XTB after 4 h at 65 °C. Samples were analyzed on a 8% native polyacrylamide gel. 1, XTB (12.5 nM) alone; 2, XTB + Int$^{SSV}$ (2 $\mu$M). C, $k_{app}$ is a linear function of the enzyme concentration.
DNA Cleavage Activity of the SSV1 Integrase

with the reverse assumption of rapid binding and slow cleavage, is not consistent with the data. According to Scheme 1, the fraction of substrate covalently bound to the enzyme, \( f_{\text{cov}} \), is equal to \( f_{\text{cov}} = E S_{\text{cov}} / S_0 \), where \( S_0 \) is the total substrate concentration, \( S_0 = S + ES + ES_{\text{cov}} = ES_{\text{cov}}(1 + 1/K + k_{-1}/k_1 E_0 K) \), since \( E_0 \gg S_0 \). \( E \) (representing the free enzyme) can be approximated to \( E_0 \), the total amount of enzyme. The reciprocal of the fraction \( f_{\text{cov}} \) of covalently bound substrate at final equilibrium is then given by, \( 1/f_{\text{cov}} = 1 + (1 + k_{-1}/k_1 E_0 K) \). Since \( f_{\text{cov}} \) was found independent of \( E_0 \), then \( k_{-1}/k_1 E_0 K \) was smaller than 1, and the ratio \( k_{-1}/k_1 \) was smaller than the lowest \( E_0 \) (0.1 \( \mu \)M), that is \( k_{-1}/k_1 < 10^{-7} \) M. In this case, the partitioning between the covalent \( ES_{\text{cov}} \) and the noncovalent ES complexes would remain independent of the total protein concentration because (almost) all the substrate would be bound (covalently or noncovalently) to the integrase, with (almost) no free substrate remaining. The technique used here could not determine the relative amount of free and noncovalently bound substrate, but we could verify on nondenaturing gels that most of the substrate was bound to \( \text{Int}^{\text{SSV}} \) as predicted by our model (Fig. 4B).

Determination of the Kinetic Parameters of the Reaction—In contrast to the final plateau, the rate at which the substrate was bound to \( \text{Int}^{\text{SSV}} \) as predicted by our model (Fig. 4A). This was also in agreement with Scheme 1, which predicts that formation of the covalent complex \( ES_{\text{cov}} \) should follow first-order kinetics, with an apparent rate constant given by the following equation.

\[
k_{\text{app}} = k_1 E_0 + k_{-1}/(1 + K)
\]  

First-order rate constants \( k_{\text{app}} \) were obtained from Fig. 4A by fitting the kinetics to exponential curves, and Fig. 4C shows that the values of \( k_{\text{app}} \) increased linearly with \( E_0 \). The rate constant for \( E_0 = 0.05 \) \( \mu \)M was obtained by assuming that the final value of \( f_{\text{cov}} \) was 0.4 as for the other values of \( E_0 \). According to Equation 1, the slope of the straight line gave \( k_1 = (2.5 \pm 1) \times 10^{-8} \) M \(^{-1}\) s \(^{-1}\) and the vertical intercept gave \( k_{-1}/k_1 = (0.8 \pm 0.6) \times 10^{-5} \) s \(^{-1}\). The ratio \( k_{-1}/k_1 \) is obtained from these kinetics was indeed lower than the limit of 10 \(^{-7} \) M deduced from the lack of variation of \( f_{\text{cov}} \) with \( E_0 \).

The rate-limiting step for the cleavage reaction would therefore be the formation of a noncovalent complex between the integrase and its substrate: high concentrations of enzyme and/or substrate would be needed to form this complex at an appreciable rate. It is a high (of the order of micromolar) concentration of (at least) one of the reagents that would be required for cleavage, rather than a high molar ratio of enzyme to substrate. Another conclusion is that after binding noncovalently to the enzyme, only about 40% of the substrate undergoes cleavage with coconant covalent binding to the protein. The equilibrium between ES and \( ES_{\text{cov}} \) (Scheme 1), that is between intact and cleaved substrate on the surface of the integrase, would have an equilibrium constant \( K \) of \( 0.4/0.6 = 0.7 \). This constant has a value near unity, consistent with an equilibrium involving an exchange between 2 phosphodiester bonds, the phosphodiester bond of the intact substrate and the phosphodiester bond with the hydroxyl group of Tyr \(^{314} \) on the cleaved substrate. An equilibrium constant near unity also indicates that the cleavage reaction could be easily reversed, thus leading to possible exchange with another DNA partner.

Affinity of the Enzyme for the Synthetic Substrates—Analysis of the relative affinity of \( \text{Int}^{\text{SSV}} \) for the XTB substrate was performed by adding the pKS-attB plasmid harboring the chromosomal attB site (8) as specific competitor into the cleavage reaction.
FIG. 6. Effect of the temperature on Int<sub>SSV</sub> activity. A, thermal stability of the enzyme. Int<sub>SSV</sub> was incubated at 82 °C without DNA. At different times, 1 μM protein samples were transferred into a test reaction containing 12.5 nM XTB and incubated for 2 h at 65 °C. The reaction products were analyzed by SDS-PAGE and the amount of covalent complex formed were determined by scanning the autoradiogram of the gel. The data were normalized by taking the amount of complexes formed without preincubation as 100%. The resulting relative activity values were plotted against the time of preincubation. B, temperature dependence of the cleavage activity. Int<sub>SSV</sub> (1 μM) was incubated for 4 h with 12.5 nM XTB substrate at six different temperatures ranging from 40 to 65 °C. The reaction products were analyzed as described in A except for normalization. The inset shows an Arrhenius plot of the data.
activity probably reflects that the overall velocity of the reaction is a function of temperature. Representation of the data in an Arrhenius plot is given in the inset of Fig. 6B. Linear approximation of the profile yielded a "break" point located around 50°C. Below 50°C, the calculated activation energy \( E_a \) is of 169.8 kJ/mol while above 50°C, it decreases to 66.5 kJ/mol. Such a profile for an Arrhenius plot can be the result of a lower than expected rate at high temperatures. A trivial explanation for such behavior is that the substrate starts to melt around 50°C, thus becoming rate-limiting above this temperature. Another possible explanation for the bimodal profile of thermodendence of the cleavage activity (inset, Fig. 6B), is that conformational transitions may occur within the enzyme in the studied temperature range, affecting either the binding of the enzyme to its substrate, or simply the equilibrium between cleavage and religation.

**Temperature-induced Changes of IntSSV Conformation**—Fluorescence measurements allow to monitor the local environment of the aromatic side chains, mostly the Trp residues (29). The 3 Trp residues present in IntSSV may therefore be probes to follow the potential conformational changes induced by temperature. Like its absorption spectrum, the fluorescence excitation spectrum of IntSSV at 20°C has a maximum at 278 nm (data not shown). However, when exciting at this wavelength the fluorescence signal is essentially due to Tyr residues (27 present in IntSSV). We therefore followed the protein fluorescence upon excitation at 295 nm at different temperatures. The fluorescence spectra emitted by IntSSV in a temperature range of 20 to 58°C are characterized by a maximal emission wavelength of 343 nm (Fig. 7A). When increasing the temperature, the fluorescence at 343 nm is reduced, to reach at 58°C about 40% of the intensity observed at 20°C. The decrease in emitted fluorescence is reversible and the full intensity recovered when going back to 20°C. The decrease in fluorescence intensity indicates a restricted flexibility of the aromatic side chains as a probable consequence of strengthened hydrophobic interactions. Tightening of the hydrophobic interactions with increasing temperature would compensate for the increased thermal agitation, thus preventing thermal unfolding. When plotting the maximal emitted fluorescence as a function of temperature (Fig. 7B) a discontinuity of fluorescence decrease is observed around 45°C. Interestingly, the temperature of the discontinuity correlates to the transition temperature in the Arrhenius plot (inset, Fig. 6B), suggesting that the break in this plot is due to conformational changes in the protein rather than to alteration of the substrate. Such a behavior has been previously reported for other thermophilic enzymes (30, 31).

**IntSSV Can Cleave Mismatched Substrates**—To know if sub-
IntSSV seems to cleave more efficiently the XT substrate than those mismatches in the binding sites do not alter substrate substrates leads to formation of a covalent complex suggesting tions equivalent to the archy of mutational effects based on the nature of the mutation of the top strand cleavage seems to be more important than its formation above 60°C and suggesting that IntSSV may recognize and cleave single-stranded substrates. However, we suspected that both the XT and XB single-stranded oligonucleotides could anneal on themselves to form imperfect double-stranded substrates that may be cleaved by IntSSV (Fig. 3C). This homopairing would generate a BspEl site, and we used this property to determine the structures of the substrates that were incubated with the integrase. Indeed, digestion of the XT substrate by BspEl prior to incubating with IntSSV totally abolished the covalent complex formation (data not shown) suggesting that the XT sequence could adopt a double-stranded structure recognized as a substrate by IntSSV. In the light of a double-stranded structure for XT, the cleavage efficiency would reflect the sum of “top” (CT in Fig. 3C) and “bottom” (CB in Fig. 3C) strand cleavages, since both strands of the substrate are 5′-end labeled in this case. If so, the XT cleavage efficiency is likely to be lower than the XTB cleavage where only one strand is labeled.

Comparison of the XT and XB homopairing shows that both produce a mismatch at position −1 of the “CT” cleavage site and a mismatch at position −2 of the “CB” cleavage site (Fig. 3C). However, incubation of IntSSV with either the XT or XB substrates leads to formation of a covalent complex suggesting that mismatches in the binding sites do not alter substrate binding and cleavage by the enzyme. As shown in Table I, IntSSV seems to cleave more efficiently the XT substrate than the XB substrate. Furthermore, formation of a covalent complex between the enzyme and the XT substrate is not affected by large amounts of nonspecific competitor (Table I: compare XT cleavage to XTB bottom strand cleavage). It should be noted that the −1 position of the CT cleavage site is identical to the XT site in the XT:XT appar- ement, but different in the case of the XB:XB apperiment (Fig. 3C). For IntSSV the chemical nature of the base in position −1 of the top strand cleavage seems to be more important than its pairing to the opposite strand, at least for the cleavage step. The influence of positions −1 and −2 on cleavage efficiency as well as substrate mismatching is now under investigation.

IntSSV is also able to bind and cut substrates harboring mismatches in the binding sites. The chemical nature of the base in position −1 of the top strand cleavage seems to be more important than its pairing to the opposite strand, at least for the cleavage step. The influence of positions −1 and −2 on cleavage efficiency as well as substrate mismatching is now under investigation.

**Acknowledgments**—We thank B. Hallet for helpful discussions and C. Bouthier de la Tour for critical reading of the manuscript.

**REFERENCES**

1. Nash, H. A. (1996) in Escherichia coli and Salmonella Cellular and Molecular Biology (Neidhardt, F. C. A. E., ed) Vol. 2, 2nd Ed., pp. 2363–2376, ASM Press, Washington, D.C.
2. Hallet, B., and Sherratt, D. (1997) FEMS Microbiol. Rev. 21, 157–178
3. Grainge, I., and Jayaram, M. (1999) Mol. Microbiol. 33, 449–456
4. Cheng, C., Kussie, P., Pavletich, N., and Shuman, S. (1998) Cell 92, 841–850
5. Sherratt, D. J., and Wigley, D. B. (1998) Cell 93, 149–152
6. Yeats, S., McWilliam, P., and Zillig, W. (1992) EMBO J. 11, 1035–1038
7. Reiter, W. D., Palm, P., and Yeats, S. (1989) Nucleic Acids Res. 17, 1907–1914
8. Muskhelishvili, G., Palm, P., and Zillig, W. (1995) Mol. Genet. Genet. 237, 334–342
9. Muskhelishvili, G. (1994) System. Appl. Microbiol. 16, 605–608
10. Palm, P., Schleper, C., Gramp, B., Yeats, S., McWilliam, P., Reiter, W. D., and Zillig, W. (1991) Virology 185, 242–250
11. Argos, P., Landy, A., Abremski, K., Egan, J. B., Haggard-Lindquist, E., Hoess, R. H., Kuhn, M. L., Kalions, B., Narayana, S. V. L., Persson, L. S., II, Sternberg, N., and Leong, J. M. (1986) EMBO J. 5, 433–440
12. Abremski, K. E. and Hoess, R. H. (1989) Proteem Eng. 5, 87–91
13. Esposito, D., and Scocca, J. (1997) Nucleic Acids Res. 25, 3605–3614
14. Nunes-Duby, S. E., Kwon, H. J., Tirumalai, R. S., Ellenberger, T., and Landy, A. (1998) Nucleic Acids Res. 26, 391–406
15. Campbell, A. M. (1992) J. Bacteriol. 174, 7495–7499
16. Casadaban, M. J., and Cohen, S. N. (1980) J. Mol. Biol. 138, 179–207
17. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
18. Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. (1995) J. Bacteriol. 177, 4121–4130
19. Masing, J. (1988) Methods Enzymol. 101, 20–78
20. Kunkel, T. A., Roberts, J. D., and Zaks, R. A. (1987) Methods Enzymol. 154, 367–382
21. Bradford, M. M. (1976) Anal. Biochem. 72, 248–254
22. Laemmli, U. K. (1970) Nature 227, 680–685
23. Schleper, C., Kubo, K., and Zillig, W. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 7645–7649
24. Pierzon, L. S., III, and Kahn, M. L. (1987) J. Mol. Biol. 196, 487–496
25. Hauser, M. A., and Scocca, J. (1992) J. Bacteriol. 174, 6859–6864
26. Dupont, L., Boizet-Bonhoure, B., Coddeville, M., Auveray, F., and Ritzenthaler, P. (1995) J. Bacterial. 177, 586–595
27. Pena, C. E. A., Stoner, J. E., and Hatfull, G. F. (1996) J. Bacteriol. 178, 5533–5536
28. Raynal, A., Tuphile, K., Gerbaud, C., Luther, T., Guerineau, M., and Pernodet, J. M. (1986) Eur. J. Biochem. 170, 325–333
29. Huang, S. (1991) J. Biol. Chem. 266, 13727–13737
30. Shuman, S. (1991) J. Biol. Chem. 266, 13727–13737