Volume 7 Number 8 1979 Nucleic Acids Research

Studies on the binding of lambda Int protein to attachment site DNA; identification of a tight-binding site in the P′ region

R.Wayne Davies*, P.H.Schreier†a, M.L.Kotewicz+b and H.Echols+

*Department of Biology, University of Essex, Colchester, CO4 3SQ, †M.R.C. Laboratory of Molecular Biology, Hills Road, Cambridge, CB2 2QH,UK, and +Department of Molecular Biology, University of California, Berkeley, California, USA

Received 11 September 1979

ABSTRACT

We have used three approaches to studying the interaction of lambda Int protein with bacteriophage attachment site DNA, POP': location of binding sites by retention of DNA fragments in a filter binding assay, reconstruction of a binding site by DNA synthesis and protection of a binding site from an exonuclease. Retention of restriction fragments on nitrocellulose filters in the presence of Int protein was used to locate binding sites. A high affinity binding site lies in P' between base pairs -6 and +173 from the center of the common core sequence, and low affinity sites are found in the 200 base pair region left of position -6. Reconstruction of the high affinity binding site region from the right using primed DNA synthesis and testing for filter binding in the presence of Int protein shows that sequences sufficient for tight binding of Int protein lie to the right of position +66. When attachment site DNA is protected by bound Int protein against digestion by exonuclease III, four Int dependent protection bands are seen in positions +58, +68, +79 and +88. This can be interpreted either as showing that four Int protein monomers bind to the high affinity region in series, or as evidence for wrapping of the DNA around Int protein, leading to structural changes resembling those occurring to DNA in nucleosomes.

INTRODUCTION

Lambda is a temperate bacteriophage that can lysogenize its host, E. coli. In the prophage state the phage achieves replicative synchrony with its host by physically integrating its DNA into the host chromosome. Integration occurs by a single reciprocal site-specific recombination event between unique sites on the phage and bacterial chromosomes known as attachment (att) sites (reviewed in refs. 1-3). Integration depends upon the product of the lambda int gene, and upon several host proteins (4, 5). Int protein has been purified and shown to bind specifically to DNA molecules containing the lambda phage attachment site (att POP') or the left prophage attachment site (att BOP') (6, 7, 8). It is not known how the host attachment site (att BOB') or the right prophage attachment site (att POB') inter-

act with the host factors and Int protein in the complete reaction.

The most likely mechanism for integrative recombination would be for Int protein to recognize and align the phage and host attachment sites, perhaps aided in the recognition of the host att BOB' by the required host factor, followed by breakage and rejoining catalysed by host proteins or possibly by Int itself (9).

The interaction of Int protein and the attachment sites is thus clearly of central importance in the recombination process. The phage attachment site has been shown to have the formal structure POP', where P and P' are flanking regions to the left and right respectively of the sequence O, and O (designated common core), represents a stretch of 15 nucleotides common to all four attachment sites (10). Although the position of common core is known exactly because it is common to all the attachment sites, the extent of sequence to the right and left which is critical in the reaction is not known. In order to determine the regions of DNA required for Int to interact with the viral attachment site we have studied the binding of purified Int protein to small DNA fragments from this region.

In particular, the location of an Alu I site one base-pair within the common core allowed us to separate the two halves of att POP' into P and OP' and to show that the main binding site is located in the OP' fragment. We have further used a new approach, protection against the action of exonuclease III, to study how far to the right in the P' sequence Int protein binding extends. In a final series of experiments presented here we reconstructed the viral attachment site from single stranded template and primer to further define the DNA sequence required for tight binding of Int. The results obtained in these studies are therefore complementary to a detailed study of the high affinity binding site in P' at the base pair level using DNA protection methods (11, 12, 13) which will be published elsewhere (S. G. Minter, R. W. Davies, M. Kotewicz, and H. Echols, manuscript in preparation).

MATERIALS AND METHODS

Preparation of DNA fragments

All DNA fragments described here were prepared from DNA of the plasmid pMG 1409 (14) which contains the λ Eco R1.C fragment (15), or from the plasmid pacl 29 constructed by G. Cesareni and S. Brenner, which contains the Hind III cut 3-Bam H1 cut 3 fragment of λ. Both of these fragments contain the phage attachment site (att POP'). Conditions for plasmid

preparation, restriction enzyme digestion, gel electrophoresis and fragment elution have been described previously (14, 16, 17). For convenience we refer to the Hind III cut 3-Bam H1 cut 3 fragment of λ as fragment α, and to the 317 base pair Hinf I fragment derived from α as β. φX174 RF DNA was a gift of F. Sanger.

End labelling of DNA fragments

Most of the experiments described here were carried out using mixtures of DNA fragments resulting from restriction enzyme digests. Since fragments labelled by nick-translation showed a higher level of non-specific binding by Int than end-labelled fragments, all fragments were end-labelled either by using polynucleotide kinase (18) or by filling in the ends generated by restriction enzyme cuts. The latter was accomplished by using the Klenow fragment of DNA polymerase I (Boehringer) to add $\alpha^{32}P$ labelled nucleotides to the recessed 3' OH ends of appropriate restriction cuts.

For example, to label Hinf I ends, 5μl of a HinfI digest of fragment α, 2.5 μl of a mix containing 0.125 mM dGTP, dTTP, dCTP, 15 mM Tris HCl pH 7.4, 15 mM $MgCl_2$, 125 mM NaCl and 2.5 mM DTT, 5 μl (dried-down) $\alpha^{32}P$ labelled dATP (5 μCi 400 Ci/m mol, Amersham), and 0.1 unit of the Klenow fragment were combined and incubated at 22°C for 30 minutes. The reaction was stopped by heating the mixture at 70°C for 10 minutes, and the labelled DNA fragments were separated from unincorporated triphosphates on a 1 ml Sephadex G-100 column equilibrated with 2.5 mM Tris-HCl, pH7.4, 0.05 mM EDTA. The samples were then concentrated by lyophilization.

Purification of Int protein

Int protein was purified as described elsewhere in detail (6, 8), and was tested for specificity of binding by comparing filter binding of in vivo labelled phage DNA carrying the λ or φ80 attachment site. Three separate preparations of Int protein were tested. The first two were greater than 85% pure as judged by SDS gel electrophoresis and the third (which was purified from an Int overproducing strain kindly provided by S. Hu and A. Honigman) although it was not as pure on protein gels was of an even higher binding specificity. They are referred to as Int1, 2, and 3 in chronological order of preparation.

Filter binding tests

End-labelled DNA fragments were resuspended in a small volume of sterile distilled water, and aliquots containing fragments equivalent to 25 nanograms of the α fragment were distributed to each assay tube. 10 X

Int binding buffer was added to give final concentrations of 10 mM Tris-HCl, pH 7.4, 10 mM $MgCl_2$, 0.2 mM EDTA, 0.2 mM DTT and 70 mM KCl, in a reaction volume of 100 to 250 µl. As the purified Int protein was always in 400 or 800 mM KCl, this was carefully noted such that the final binding conditions did not exceed 70 mM KCl. After 5 minutes at 0°C, various amounts of Int were added, and incubation continued for 10 minutes at 0°C. The solution was filtered through a millipore filter (0.45 µm, type HA) at a flow rate of 1 ml per minute. Filters were presoaked for at least 30 minutes in binding buffer containing 70 mM KCl. All filtration was done at 22°C. To remove weakly bound fragments, the filter was washed 2X with 1 ml of binding buffer (70 mM KCl), 1X with 1ml binding buffer (0.5 M KCl), and 1X with 1 ml of binding buffer (70mM KCl). Each wash was collected directly into centrifuge tubes. Strongly bound DNA was eluted with 1 ml of 0.5% SDS. To each filtrate 25 µg of E. coli tRNA, 0.1 volume of 3.0 M Na-acetate, and 3.0 volumes of ethanol were added and mixed. Any SDS precipitate that formed was removed by gentle centrifugation. The solutions were chilled and centrifuged (30' at 35000 rpm, SW 50l Beckman). The precipitates were washed twice with prechilled 95% ethanol and dried. Recovery was always 90% or greater for these precipitations. The dried precipitates were resuspended in 5 µl of H_2O and 2.5 µl of a solution of 15% Ficoll containing 0.03% bromphenol blue and 0.03% xylene cyanol was added. The samples were electrophoresed in 8% polyacrylamide thin gels (18). After autoradiography, fragments were cut out, and counted in 10 ml of Aquasol (New England Nuclear).

Protection against exonuclease III digestion

DNA substrate was prepared by primed incorporation on a purified λ r-strand as described previously (16). Approximately 40 ng of the 55 base-paid MboII-Hinf I fragment was mixed with 2 µl of purified λ r-strand (1.0 mg/ml) and 0.5 µl of 10X Hin buffer (66 mM Tris-HCl, pH 7.4, 66 mM $MgCl_2$, 500 mM NaCl, and 10 mM DTT). After annealing the primer to the template, 10 µl of H_2O, 4 µl of a mix containing 0.125 mM dCTP, dTTP, dGTP and 2.5X Hin buffer were added. This mixture was added to 2 µCi of dried α^{32}p labelled dATP (400 Ci/mmol Amersham), 0.2 units of Klenow fragment of DNA polymerase I was added (Boehringer) and the solution was incubated for 10 minutes at 22°C. The reaction was stopped by adding 2.5 µl of 0.25 M EDTA and heating 10 minutes at 70°C. The DNA was purified by passing it over a G100 column. Each experiment contained a molar equivalent of 2.5 ng. of fragment α. The DNA was then incubated with various amounts of Int protein

in Int binding buffer at 0°C as before, except that the DTT concentration was raised to 1.0 mM. Exonuclease III (New England Biolabs) was then added (4 to 24 units), and the mixture incubated at 22°C or 37°C for 5 to 15 minutes. Under the conditions used for the experiments shown in figure 5, the labelled strand was completely digested in the absence of Int protein. The Exonuclease III digestion was stopped by the addition of 5 µl of 0.25 M EDTA, the DNA precipitated in the presence of 50 µg/ml E. coli tRNA, washed with prechilled 95% ethanol and resuspended in sterile distilled H_2O. An equal volume of deionised formamide (98%) containing 0.03% bromphenol blue and 0.03% xylene cyanol was added, the sample boiled for 3 minutes, and electrophoresed in a thin 7 M urea, 8% polyacrylamide gel. Samples from sequencing experiments using the chain termination procedure (20) were run as length standards.

Reconstruction of the Int binding site

DNA substrates having a common right end and a variable left end were produced by priming pure λ 1-strand template with the 28 base pair Mbo II-Hinf I fragment (Fig 6). If this primer is extended to the left through the Int binding sequences the fragment will be retained on nitrocellulose filters in the presence of Int protein. In order to avoid artifactual binding of the substrate, single-stranded regions were removed with S1 nuclease.

Priming reactions were carried out as described above, except that aliquots were taken at 30 second intervals in order to generate a series of substrates extending various distances to the left. This priming produced highly labelled substrates. Reactions were stopped with 25 µl of phenol, extracted 6 times with ether, purified over a Sephadex G-100 column as before, and lyophilised. The DNA was resuspended in 20 µl of S1 buffer (30mM Na acetate, pH 4.6, 50 mM NaCl, 1 mM $ZnSO_4$), and digested with S1 for 30 minutes at 22°C. The conditions of S1 digestion were determined empirically for each DNA preparation by adding various amounts of S1 to small portions of the DNA to be tested. Concentrations of S1 were chosen where the DNA just lost its capacity to bind to the filters, i.e. where it had lost its single stranded regions.

Using the predetermined S1 conditions, DNA was treated as above. Aliquots containing the molar equivalent of 25 ng of fragment α were incubated with Int and assayed for filter binding as described above. The filtrates were ethanol precipitated, boiled in formamide, and electrophoresed under denaturing conditions as described for the exonuclease III

experiments.

RESULTS

The DNA sequence of the phage attachment site (att POP') and some of the surrounding region has been determined (10, 16), so that the locations of recognition sites for restriction enzymes are known very precisely. Figure 1 shows the positions of the sites within the ∝ fragment of λ. The Alu I recognition sequence is found in the phage att site (att POP'), and the enzyme cuts one base pair into the left end of the common core sequence.

In preliminary experiments (data not shown) we found that Int protein purified and tested as described by Kotewicz et al (6) bound the λ Eco RI fragment C (15) and the λ Hind III_3-Bam HI_3 fragment (i.e. fragment ∝, 14) both of which are very likely to contain the entire phage att site (7). The ∝ fragment was digested with Hinf I and the resulting fragments were end-labelled. Nick translation was not satisfactory for labelling as it gave higher backgrounds of non-specific binding. Using the mixture of fragments within each filter binding experiment allowed direct comparison of the ability of Int protein to bind each fragment. The labelled restriction fragments were incubated with increasing amounts of Int protein, passed

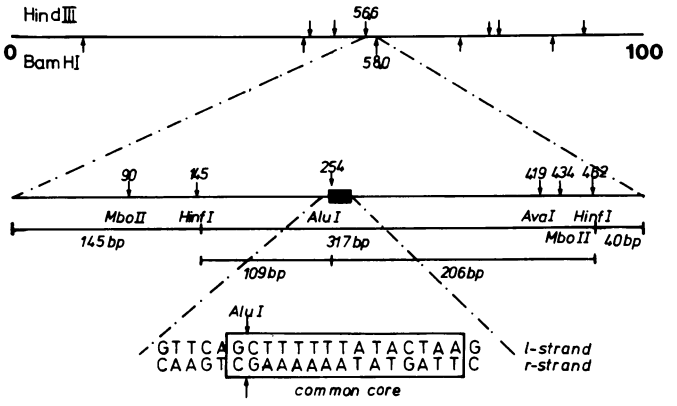


Figure 1. The locations of the restriction cuts mentioned in this work. The top line represents the λ chromosome, Hind III cuts being shown above the line, Bam Hl cuts below the line. The second line shows the region from Hind III cut 3 to Bam Hl cut 3 (fragment ∝) in expanded form. Below this the DNA sequence of the common core and a few bases to the left (15, 20) is given showing the precise position of the AluI cut. The 317 bp Hinf I fragment is referred to as fragment β in the text.

through a millipore filter, and then washed as described in Materials and Methods. Figure 2A shows the elution pattern of the Hinf I subfragments of α during the washing procedure. In the absence of Int, all fragments were quantitatively recovered in the flow through and the first wash. In the presence of Int protein, all three Hinf I subfragments are retained through the low salt (70mM KCl) wash. All preparations of Int tested retained non-att DNA fragments (including fragments from φX174 and from other parts of the lambda genome, figure 2B) through this low salt wash.

When the KCl concentration in the wash buffer was raised to 0.5 M, all of the 40 base pair fragment, most of the 145 base pair fragment, but only a small amount of the 317 base pair fragment of α washed through. DNA fragments from φX174 or other regions of λ were invariably released from the filters under these conditions (figure 2B). Increasing the KCl concentration to 1.5 M did not release more DNA. This shows that the complex between Int and the 317 base pair fragment is very tightly bound. Throughout this paper, retention on a filter in the presence of a high salt wash (0.5 M KCl) will be referred to as tight binding.

The tight binding complex was eluted from the filter when SDS (0.5%) was included in the wash.

As noted above, most of the 145 base pair fragment was eluted from the filter with the high salt wash.

Three fragments, the 145 base pair fragment, a 280 base pair Hind III-Hinf I fragment that included the 145 fragment (data not shown), and a 109 base pair Hinf I-Alu I fragment that is adjacent and to the right of the 145 base pair fragment (Figures 1, 3) were partially bound at high salt if a large amount of Int was added (figure 2C). No φX174 fragments, nor other λ pieces tested bound under these conditions (figure 2B). This is apparently a specific but weak interaction. With a preparation of Int from a plasmid-containing overproducer (which gave a much higher specific activity of Int binding) this weak binding was still visible, but was further reduced. This may suggest that there is a long region of weak specific binding to the left of the common core, starting within the Hinf I-Alu I fragment and extending at least 50 base pairs into the 145 base pair fragment (Figure 1).

In the next series of experiments the enzyme Alu I was used to dissect the phage attachment site. Alu I cuts the 317 base pair fragment into a 109 base pair fragment with one base pair of the common core and a 211 base pair fragment with the other 14 base pairs of the common core (Figure 1; the longest strand is given in each case). The larger fragment is thus

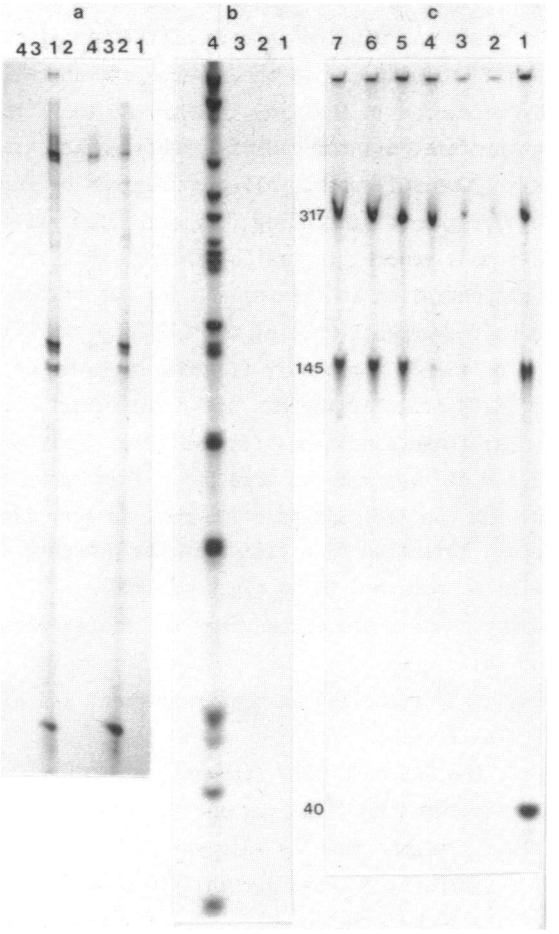


Figure 2. a. Autoradiograph of an 8% polyacrylamide gel of ethanol precipitated wash fractions from a filter binding experiment in which 25 µl of Int 1 were incubated with 50 ng of Hinf I digested kinase end-labelled fragment «, filtered and washed as in Materials and Methods. The right hand four tracks are from the incubation with Int, the left hand four without Int. Tracks labelled 1 = 2 ml binding buffer, 70 mM KCl wash, tracks labelled 2 = binding buffer, 0.5 M KCl wash, tracks labelled 3 = binding buffer, 1.5 M KCl wash, tracks labelled 4 are the 0.5% SDS wash.

b. Autoradiograph of an 8% polyacrylamide gel of ethanol precipitated 0.5% SDS wash fractions from a filter binding experiment in which various amounts of Int 2 were incubated with 200 ng of fill-in end-labelled Hinf I cut ϕX174 RF. Tracks 1, 2 and 3 had 0, 5 and 10 µl of Int 2 respectively. Track 4 shows the DNA input.

c. Autoradiograph of an 8% polyacrylamide gel of ethanol precipitated 0.5% SDS wash fractions from an experiment in which 25 ng aliquots of Hinf I cut, fill-in end-labelled fragment « were incubated with various amounts of Int 1. Tracks 1 to 7 are: DNA input, 0, 0.2, 1, 5, 10 and 25 µl Int 1.

essentially OP' with a one base pair truncated core, the small fragment is P plus one base pair of the core.

Figure 4 shows an autoradiogram of a gel of the 0.5% SDS wash fractions from a filter binding experiment in which increasing amounts of Int protein were incubated with a partial Alu I, complete Hinf I digestion of fragment ∝. The ∝ fragment was cut with Hinf I and Alu I in order to allow a direct comparison of the binding of the Hinf I-Alu I fragments with the binding of the 145 and 40 base pair Hinf I fragments whose behaviour had been previously studied. In order to have some fragment β present as an internal control for tight binding, ∝ was only partially digested with AluI. Figures 3 and 4 show that the 211 base pair Alu I-Hinf I fragment is bound to

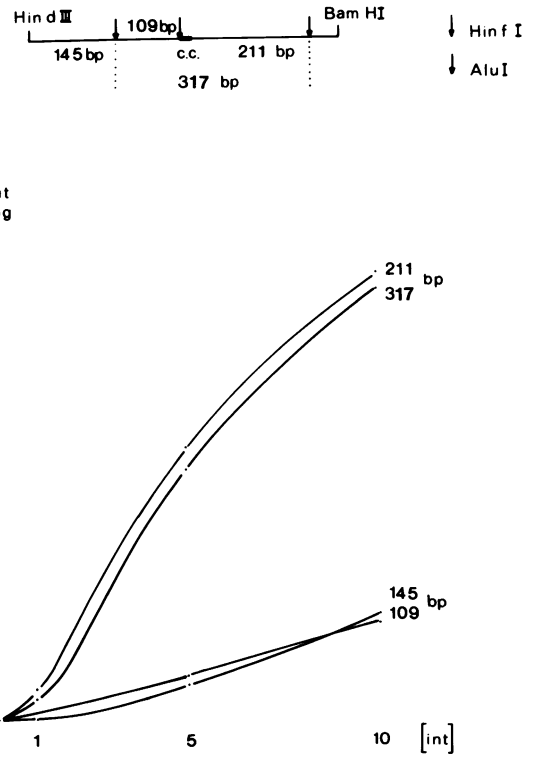


Figure 3. Graphical representation of the percentage of each of four fragments specifically bound to nitrocellulose filters in the presence of increasing amounts of Int 2. The abscissa gives µl of Int 2 added to 100 µl incubations containing 25 ng of digested fragment ∝. The experimental procedure is described in Materials and Methods. The identity of the fragments can be seen from Figure 1 and the text.

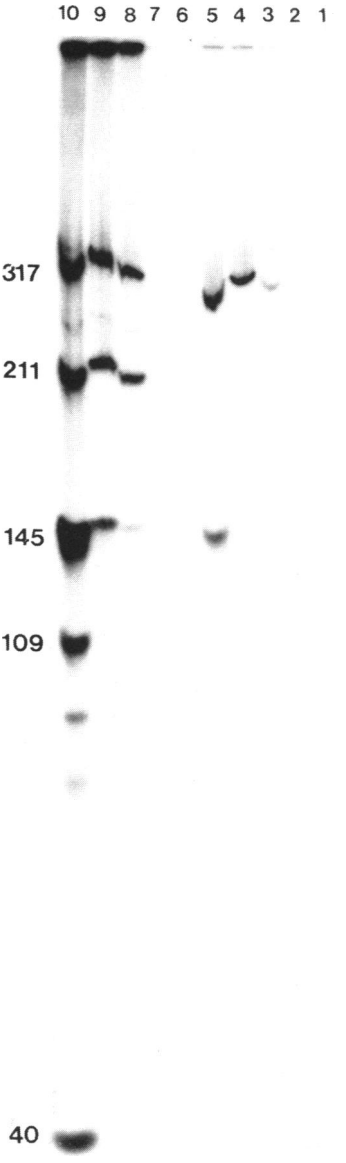


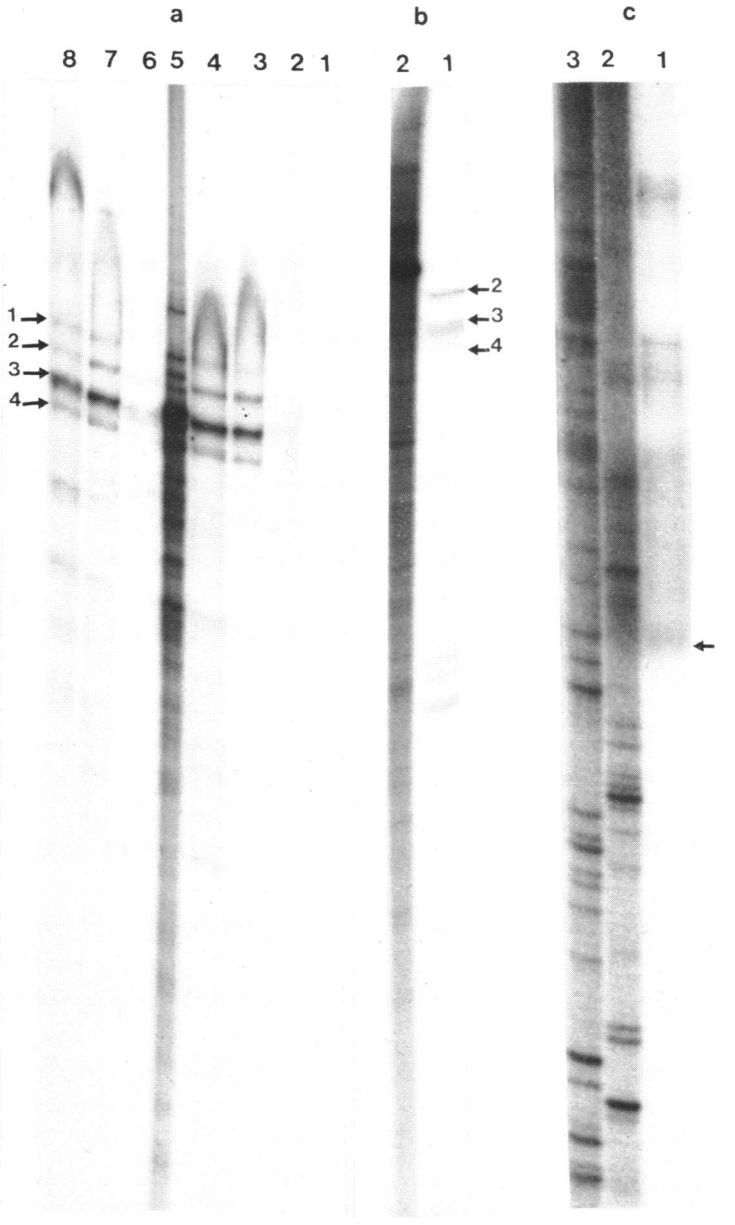
Figure 4. Autoradiograph of an 8% polyacrylamide gel of ethanol precipitated 0.5% SDS wash fractions from an experiment in which 0, (tracks 1 and 6), 1 µl (2 and 7), 5 µl (3 and 8) or 10 µl (4 and 9) Int 2 were incubated as described in Materials and Methods with 25 ng of fragment « digested either completely with Hinf I alone (tracks 1-5) or completely with Hinf I and partially with Alu I before end-labelling by filling-in. Tracks 5 and 10 show the DNA input.

nitrocellulose filters by Int protein to the same extent as is fragment β. Removing 109 base pairs of β, the P sequences and one base pair of the core, has not affected the binding of Int protein at all. Cutting off 28 base-pairs from the right end of β with MboII also did not change the binding properties of the remaining fragment (data not shown). Thus DNA sequences sufficient to account for all the observed tight binding by Int protein must lie between the second base pair of the common core and the MboII cut just before the right end of fragment β.

We undertook two further series of experiments to obtain a more precise position for the highest affinity binding site for Int. Exonuclease III digests one strand of a duplex DNA molecule from an exposed 3' OH end; such a structure is produced following incorporation of nucleotides using a small primer fragment hybridised to a single-stranded template (Figure 6). We used the 145 base pair Hinf I fragment or the 55 base pair Mbo II-Hinf I fragment as primers on a purified λ r strand template. The Klenow fragment of DNA polymerase I was used to synthesise radioactively labelled DNA through the attachment site. The products of the reaction were purified and exposed to the action of exonuclease III in the presence and absence of Int protein. If the DNA is now denatured and run in a 7M urea-polyacrylamide gel alongside a standard DNA-sequence it is possible to determine at which nucleotide(s) Int protein begins to protect against the action of exonuclease III. The 1-strand extensions that have gone beyond the Int binding site will be digested back to the point where bound Int protein interferes with further digestion by exonuclease III. This will lead to an accumulation of fragments which are seen on an autoradiogram as bands of a length corresponding to the distance from the beginning of the primer to the right end of the protected region (figure 5). The length of the DNA fragments in the bands can be determined by comparison with length standards from a sequencing experiment so that the position of the accumulation can be pinpointed in the sequence within several nucleotides. In order to relate the lengths seen on the gels to the numbering reference system of nucleotides that has its zero point at the eighth base pair of the common core (10), we report fragment lengths in base-pairs from the 1-strand Hinf I cut at the left end of β, and in brackets refer to the core numbering system. Thus base-pair 116 is base-pair 0, and base-pair 175 for example is base-pair +59.

In a series of experiments with exonuclease III, bands corresponding to accumulation of 1-strand ends due to Int protection occurred in a

region 58 to 88 base pairs to the right of the common core zero. In figures 5A and 5B a pattern of bands can be seen corresponding to positions 205 (+89), 196 (+80), 185 (+69) (the strongest band), and 175 (+59). Such



bands do not occur if the DNA used does not contain att POP' sequences; in this case complete digestion occurs even in the presence of Int protein. The ratio of Int protein to DNA in these experiments corresponded to the specific binding range as determined by the filter binding of restriction fragments.

It is striking that not just one band is found. The presence of four bands suggests a more complex interaction between Int and the P' sequences (see discussion).

Another approach to studying the interaction of Int with its binding site(s) is to reconstruct the site(s) by extending primed DNA synthesis through the attachment site region, stopping synthesis at a series of time points in order to produce a staggered distribution of endpoints in this region. In this way an array of fragments is produced that must contain both fully and partially reconstructed binding sites. Which members of the array of fragments contain sequence information necessary and sufficient for Int binding can now be studied by Int-directed filter binding after first removing the single-strands by the action of the single-strand specific nuclease S1.

The DNA was incubated with Int protein, bound to filters, washed, extracted, and run on a denaturing gel alongside a DNA length standard. It is expected that only those fragments containing sufficient sequence information to reconstitute a functional Int protein binding site will be

Figure 5. a. and b. Autoradiographs of 7M urea 8% polyacrylamide gels of exonuclease III treated Int-DNA complexes. Experiments were carried out as in Materials and Methods, Int and DNA being mixed at 0°C, brought to digestion temperature and exoIII added. In 5a 1 µl exonuclease III was used (8 units), and in 5b 3 µl (24 units) were used. In tracks a 1, 2, 3 and 4, 0, 1 5 or 10 µl Int 2 were used respectively, exonuclease III digestion being for 15 minutes at 22°C. Track a 5 is a human mitochondrial DNA sequence standard. Track a 6; 1 µl Int 2 added at the beginning, and at 5 and 10 minutes during a 15 minute digestion with exonuclease III at 22°C. Track 7 is the same as track 6, but 5 µl of Int 2 was added at each time. Track 8; 5 µl of Int 2 added at the start, and after 1 and 2½ minutes of a 5 minute exonuclease III digestion at 37°C. In Figure 5b track 2 is a standard DNA sequence, and in the sample in track 1 conditions as in 5a track 2 were used, except that the DNA was subsequently purified over a Sephadex G-100 column to prevent the salt effects seen at the top of 5a. Arrows indicate the Int-specific bands.
c. Autoradiograph of a 7M urea 8% polyacrylamide gel showing in track 1 the 0.5% SDS wash fraction of a filter binding experiment in which the DNA substrate was made by priming DNA synthesis across the attachment site with the 28 base pair Mbo II-Hinf I fragment and subsequently removing the single stranded region with nuclease S1 as described in Materials and Methods. Tracks 2 and 3 are human mitochondrial DNA sequence standards. The arrow shows the first visible band.

retained on the filter. Thus the first bands to appear in the gel correspond to DNA sequences just beyond those necessary and sufficient for retention on the filter. The results are shown in figure 5C. We used a 28 base pair Mbo I-Hinf I fragment of β (figure 6) as a primer on a λ 1-strand template, extending its 3' end towards the common core. No fragment shorter than 150 base pairs is bound to the filter (see discussion).

These experiments corroborate and extend the results of the restriction binding experiments.

DISCUSSION

We have used three approaches to study the interaction betwen the Int protein and DNA sequences in the bacteriophage lambda attachment site, POP'.

Filter binding experiments show that a high affinity binding site for Int protein lies between base pair 110 and 289 from the Hinf I cut at the left end of fragment β or between -6 and +173 from the center of the common core. All restriction fragments containing this region are bound tightly to nitrocellulose filters in the presence of Int protein, remaining stably bound in up to 1.5M KCl. Although Int protein may interact with other regions of DNA during the recombination reaction, the fact that this region is necessary and sufficient for maximal binding of Int protein implicates it strongly in the initiation of the recombination complex.

In experiments in which the attachment site region was reconstructed by primed DNA synthesis on a single-stranded DNA template from right to left (Fig. 6B), all double-stranded fragments containing at least sequences to the right of base pair 133 (+67) were bound to nitrocellulose filters in the presence of Int protein. Fragments containing only DNA to the right of this point were not bound detectably. Clearly the region immediately to the right of base pair +66 contains a high affinity binding site for Int protein. It is worth noting that this corresponds in part to a sequence noted by Davies et al. (16) to resemble the recognition site sequence for RNA polymerase, and contains a 10 base pair sequence that occurs as part of an inverted repeat structure at the ends of Tn 10 (30).

Results of experiments in which binding of Int protein protected attachment site DNA against digestion by exonuclease III also place a high affinity binding site for Int protein between base pairs +50 and +89. This is consistent with the results of the reconstruction experiments and sets a rightward limit to the binding site. If we assume that protection represents actual obstruction of the passage of exonuclease III by the bound Int

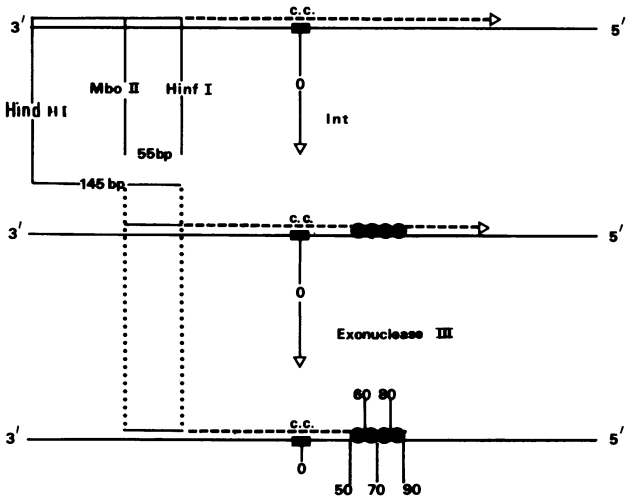


Fig.:6A

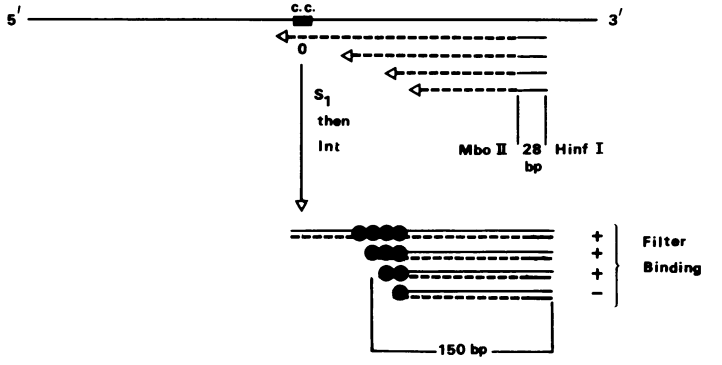


Fig.: 6B

Figure 6. Diagrammatic representation of the experiments involving protection against exonuclease III digestion (6A) or reconstruction of the high affinity binding site by primed synthesis (6B). Int is shown as binding in a four monomer series for the sake of illustration; we do not imply preference for this model.

protein, we can place the outermost limit of the region that is in contact with Int protein at base pair 205 (+89). This agrees well with the results of footprinting (11) experiments (Minter, S., Davies, R. W., Kotewicz,

M. L. and Echols, H., manuscript in preparation), which place the boundaries of the protected region corresponding to the high affinity binding site at base pairs 165 (+49) and 204 (+88). Similar experiments were reported by Ross et al (31), who place the boundaries of the protected region at base pairs 169 (+53) and 204 (+88).

The simplest model of Int-DNA-exonuclease III interaction predicts a single position at which exonuclease III should be obstructed, with at most a variability of a base-pair or two. Instead of this, we find repeatably a pattern of 4 bands at positions corresponding to base-pairs 205 (+89), 196 (+80), 185 (+69) - the strongest band, and 175 (+59), as shown in Figures 5 and 6A. We propose two possible models to explain this pattern of 4 bands spaced approximately 10 base-pairs apart.

One possibility is that there are four contiguous binding sites for Int protein in P'. Each obstruction point marks the edge of a region protected by Int. If this is so, each Int molecule must cover at most a region of 10 base-pairs. For T4 gene 32 protein, which has a monomer molecular weight of 32,000, it has been calculated (21) that the monomer covers approximately 10 base pairs. Thus it is conceivable that Int protein, which has a molecular weight of 40,000 (8), could bind as few as 10 base pairs. There are precedents for arrays of multiple binding sites, in the cases of lambda repressor (22, 23) and SV40 T antigen (24). Inspection of the DNA sequence does not reveal any precisely repeating unit but the sequence from +74 to +83 that as mentioned above also occurs at the ends of Tn10 (30) could be regarded as resembling two base pair sequence blocks to the left. The Tn10 sequence is 5' AAAATCATTA 3'; to the left of it, separated by one base pair, is the sequence TCAGTCAAAA, which has 5 out of 10 bases identical, followed immediately by the sequence CAGGTCACTA with 6 out of 10 bases identical, and all of these are within the high affinity binding site. Moreover this mode of binding could well involve cooperative interactions between Int monomers such that once the highest affinity site is occupied by an Int monomer the binding of other monomers to relatively weak sites is facilitated, as in the case of lambda repressor (Ptashne, M., personal communication). From the relative strength of the bands it would seem that the DNA sequence from 175 (+59) to 185 (+69) is bound most strongly. This model is used for purposes of illustration in Figure 6B.

Another possible explanation of these findings is that Int protein interacts with part or all of the +58 to +88 region in a way reminiscent of

the interaction of histones with DNA. It has been suggested (25) that only one side of the double helix is in contact with these proteins, resulting in the DNA being wrapped around the histone complex. This places the DNA under structural constraints which can be detected by the occurrence of typical ordered cleavage patterns on endonuclease digestion (26, 27, 28). Particularly relevant are the experiments of Riley and Weintraub (29), who showed that digestion of nucleosomes with exonuclease III results in DNA fragments displaying discrete single-stranded lengths that are multiples of 10 bases. On this model the strong bands seen as a result of protection of the DNA by Int are the result of exonuclease III being forced to pause at some repeating structural feature of the DNA, as in nucleosomes. Riley and Weintraub also found that exonuclease III can follow a strand of DNA into the nucleosome with no apparent disruption of the nucleosome's structure, so that 10 base pair repeat structures are generated stepwise, and presumably delineate only the extent of the region of structural perturbation caused by the DNA being wrapped around the protein(s). If the DNA really is altered in structure by Int binding, this has far-reaching consequences for the mechanism of recombination. This model receives support from the results of partial DNase I digestion (Minter, S., Davies, R. W., Kotewicz, M. and Echols, H., manuscript in preparation), but on the evidence presented here there is no reason to prefer it.

High amounts of Int bound weakly to the 109 bp fragment containing one base pair of the common core, the 145 bp P fragment adjacent to this, and a 280 bp fragment containing the 145 bp fragment and extending to the next Hind II site left of the common core. This binding was weak but was clearly greater than background nonspecific binding to other lambda restriction fragments or ϕX174 fragments. Thus weak binding sites for integrase probably exist left of common core, one at least being to the left of -116.

Our experiments indicate a special importance for the P' region of the attachment site in the integration reaction. A strong preference of Int protein for attachment sites containing the P' sequence has been noted previously (7). The role of the common core in Int binding is not clarified by these experiments. There cannot be a high affinity binding site spanning common core because cutting off the first base of the core with Alu I does not affect the level of filter binding. It is still possible that part or all of common core is involved in an additional lower affinity binding site.

ACKNOWLEDGEMENTS

This work was supported by a Medical Research Council project grant to R.W.D., by the MRC through F. Sanger, by the Deutsche Forschungsgemeinshaft by a grant through SFB 74 to B. Müller-Hill, by an EMBO short term fellowship to P.H.S., by a Jane Coffin Childs Fellowship to M.L.K. and by U.S. Public Health Service Grant GM 17078 to H.E. We would like to thank Dr. F. Sanger and Dr. J. C. Wang for encouragement and for providing laboratory facilities, and Drs. G. Cesareni and S. Brenner for the plasmid pacl 29.

[a]Present address: Institut für Genetik der Universität zu Köln, 5 Köln 41, Weyertal 121, GFR

[b]Present address: Department of Biochemistry and Molecular Biology, Harvard University, Biological Laboratories,16 Divinity Avenue, Cambridge, MA 02138, USA

REFERENCES

1 Gottesman, M.E. and Weisberg, R.A. (1971) in The Bacteriophage Lambda (ed. Hershey, A.D.), Cold Spring Harbor Laboratory, New York, 329-354.
2 Nash, H.A. (1977) Curr. Topics Microbiol. Immunol. 78, 171-199.
3 Weisberg, R.A., Gottesman, S. and Gottesman, M.E. (1977) in Comprehensive Virology, 8 (ed. Fraenkel-Conrat, H. and Wagner, R.) New York, Plenum Press, 197-258.
4 Williams, J.G.K., Wulff, D.L. and Nash, H.A. (1977) in "DNA Insertion Elements, Plasmids, and Episomes" (ed. Bukhari, A.I., Shapiro, J.A. and Adhya, S.L.) Cold Spring Harbor Laboratory, New York, 357-361.
5 Miller, H.I. and Friedman, D.I. (1977) in "DNA Insertion Elements, Plasmids, and Episomes" (ed. Bukhari, A.I. Shapiro, J.A. and Adhya. S.L.) Cold Spring Harbor Laboratory, New York, 349-356.
6 Kotewicz, M.L., Chung, S., Takeda, Y. and Echols, H. (1977) Proc. Nat. Acad. Sci. U.S.A. 74, 1511-1515.
7 Kikuchi, Y. and Nash, H.A. (1978) J. Biol. Chem. 253, 7149-7157.
8 Kotewicz, M.L., Grzesiak, E., Courschesne, W., Fischer, R., and Echols, H. (1979) J. Biol. Chem., submitted.
9 Nash, H.A. (1979) Proc. Nat. Acad. Sci., submitted.
10 Landy, A. and Ross, W. (1977) Science 197, 1147-1160.
11 Galas, D.J. and Schmitz, A. (1978) Nucleic Acids Res. 5, 3157-3169.
12 Gilbert, W., Maxam, A.M. and Mirzabekov, A.D. (1971) in Control of Ribosome Synthesis, Alfred Benson Symposium IX, Kjeldgaard, N. and Maaloe, O., Eds. pp 139-148, Munksgaard, Copenhagen.
13 Gilbert, W. (1972) in Polymerisation in Biological Systems, pp 245-259, Elsevier, North-Holland, Amsterdam.
14 Schreier, P.H., Davies, R.W., Büchel, D.E., Gronenborn, B., Fanning, T.G., von Wilcken, B. and Messing, J. (1977) Nature 267, 555-557.
15 Thomas, M. and Davis, R.W. (1975) J. Mol. Biol. 91, 315-328.
16 Davies, R.W., Schreier, P.H. and Büchel, D.E. (1977) Nature 270, 757-760.
17 Davies, R.W., Schreier, P.H. and Büchel, D.E. (1978) Nucleic Acids Res. 5, 3209-3218.

18 Maxam, A.M. and Gilbert, W. (1977) Proc. Nat. Acad. Sci. U.S.A. 74, 560-564.
19 Sanger, F. and Coulson, A.R. (1978) FEBS Lett. 87, 107.
20 Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Nat. Acad. Sci. U.S.A. 74, 5463-5467.
21 Alberts, B.M., and Frey, L. (1970) Nature 227, 1313-1318.
22 Ptashne, M., Backman, K., Humayun, M.Z., Jeffrey, A., Maurer, R., Meyer, B. and Sauer, R.T. (1976) Science 194, 156-161.
23 Humayun, M.Z., Kleid, D. and Ptashne, M. (1977) Nucleic Acids Res. 4, 1595-1607.
24 Tijan, R. (1978) Cell 13, 165-179.
25 Mirzabekov, A.D. and Rich, A. (1979) Proc. Nat. Acad. Sci. U.S.A. 76, 1118-1121.
26 Noll, M. (1974) Nature 251, 249-251.
27 Sollner-Webb, B. and Felsenfeld, G. (1977) Cell 10, 537-543.
28 Whitlock, J.P., Rushizky, G.W. and Simpson, R.T. (1977) J. Biol. Chem. 252, 3003-3006.
29 Riley, D. and Weintraub, H. (1978) Cell 13, 281-294.
30 Kleckner, N. (1979) Cell 16, 711-720.
31 Ross, W., Landy, A., Kikuchi, Y and Nash, H.A. (1979) Cell 18, 297-308.