JOURNAL OF VIROLOGY, July 1994, p. 4220–4226
0022-538X/94/$04.00+0
Copyright © 1994, American Society for Microbiology

Vol. 68, No. 7

Regulation of *int* Gene Expression in Bacteriophage P2

A. YU,† V. BARREIRO, AND E. HAGGÅRD-LJUNGQUIST*

Department of Cell and Molecular Biology, Karolinska Institutet, S-171 77 Stockholm, and
Department of Genetics, Stockholm University, S-106 91 Stockholm, Sweden

Received 3 February 1994/Accepted 9 April 1994

The *int* gene of bacteriophage P2 is the only viral gene necessary for the integration of P2 into the *Escherichia coli* host chromosome. This gene is situated between the phage attachment site, *att*P, and the repressor *C* gene, and is cotranscribed with *C* from the P_c promoter, towards *att*P. The P_c promoter is negatively controlled by the *cox* gene, which is the first gene of the early operon. In vitro recombination assays have indicated that in P2 an overproduction of Int is deleterious to the integrative process. We report here that the level of *int* expression is affected by several different mechanisms after transcriptional initiation. First, a partial transcription termination signal located between the *int* and *C* genes reduces the transcriptional readthrough by about 30%. Second, the ribosome binding site and AUG codon of the *int* gene are located in a putative stem-loop structure, which may inhibit the initiation of translation. The *nip1* mutation (a G to A substitution at the 22nd coding nucleotide of *int* which results in an increased efficiency of excision) is shown to relieve this inhibition, possibly through the formation of an alternative mRNA secondary structure. However, the third and probably most important control of *int* expression in P2 seems to be that of posttranscriptional autoregulation. The binding site of the Int protein on *int* gene mRNA is shown to extend into the ribosome binding site of *int*, supporting our earlier proposed model of competitive binding between Int and ribosomes.

In the site-specific recombination system of P2, integration is mediated solely by the phage Int protein and the integration host factor (IHF) of *Escherichia coli*. Both in vivo and in vitro, this is an efficient process involving four components, *att*P, Int, IHF, and *att*B, the first three of which are believed to form a complex nucleoprotein structure, the intasome, prior to synapsis, strand exchange, and ligation (37). On the other hand, analogously to λ (17), *att*B is believed to enter the reaction as non-protein-bound DNA, and the *att*P intasome is thought to slide along the chromosome in search of its reactive partner. The strong dependency on Int concentration for optimal *att*P × *att*B recombination in vitro (38) suggests that in vivo the expression of Int must be tightly regulated for P2 to integrate successfully.

Excision of the prophage requires the *cox* gene product in addition to Int and IHF. A Cox binding site on the right arm of *att*P has recently been reported, and it is likely that the Cox protein has a function analogous to that of Xis in λ, i.e., that by interaction with the other two protein components, Cox facilitates excision but inhibits integration (38). In contrast to λ, P2 does not excise easily. When a P2 lysogen with a temperature-sensitive repressor is derepressed at high temperature, at most 1% of the bacteria will produce phage (3). This appears to be because of insufficient *int* gene product, as induction approaches 100% if the heat-treated bacteria also carry a multicopy plasmid containing an active *int* gene (18). This is the direct consequence of the fact that the *int* and *cox* genes are transcribed on two divergent, mutually exclusive transcriptional units (26) (Fig. 1A). This arrangement of the genes of the recombinational proteins around *att*P is similar to that of phage 186 (7) but different from that of phage λ (9). The paradox is that lysogens of wild-type P2 have a measurable spontaneous phage production (2), because in order to attain this, P2 must express both *int* and *cox* simultaneously. The hypothesis of an alternative, *int*-specific promoter which directs the expression of *int* but not *C* has yet to be tested.

The regulatory mechanisms governing *int* gene expression in P2 are unknown, but it is evident that precise levels of Int are essential to the integrative and excisive processes in vivo. We have therefore looked into the regulation of *int* and report that this can occur at several levels, one of which is posttranscriptional autoregulation. We show that the Int protein binds to mRNA in vitro, and we locate the Int binding site on its own transcript.

MATERIALS AND METHODS

Bacterial strains. All strains are *E. coli* derivatives. C-1a, a prototrophic C strain (28), was used for all cloning procedures, *cat* activity measurements, and P2 infections. The *galK* activity measurements were done with *E. coli* C600K$^-$ (*galE*$^+$ *T*$^+$ *K*$^-$ *lac*$^-$ *thr*$^-$ *leu*$^-$) (20). The *lacZ* activity measurements were done with derivatives of strain JM109(DE3) (35).

Construction of plasmids containing progressive unidirectional deletions of the *C-int* spacer region. The *Hpa*I-*Sna*BI (Fig. 1) fragment of P2 containing the spacer region and *int* gene was inserted into the *Sma*I site of pUC18 (23). Subsequent cleavage of the recombinant plasmid with *Sph*I and *Xba*I resulted in a 4-base 3′ protruding end and a 5′ protruding end. Exonuclease III (Promega) was used to specifically digest DNA from the 5′ protruding end, and reactions were terminated at various time intervals. After religation and amplification by transformation, plasmids containing deletions of interest were identified by sequencing. The *Hin*dIII-*Eco*RI fragment from these plasmids, containing *int* and the deletions specified in Fig. 1, was inserted into the *Sma*I site of pMG524 (14) in such a way that the expression of *int* was directed by λP_L.

Plasmids. All in vitro constructions were performed by standard procedures (27). Unless otherwise stated, the en-

---

* Corresponding author. Mailing address: Department of Genetics, Stockholm University, S-106 91 Stockholm, Sweden. Phone: (46 8) 16 12 70. Fax: (46 8) 16 43 15. Electronic mail address: Elisabeth.Haggard@genetics.su.se.

† Present address: Department of Surgery, St. Göran's Hospital, S-112 81 Stockholm, Sweden.

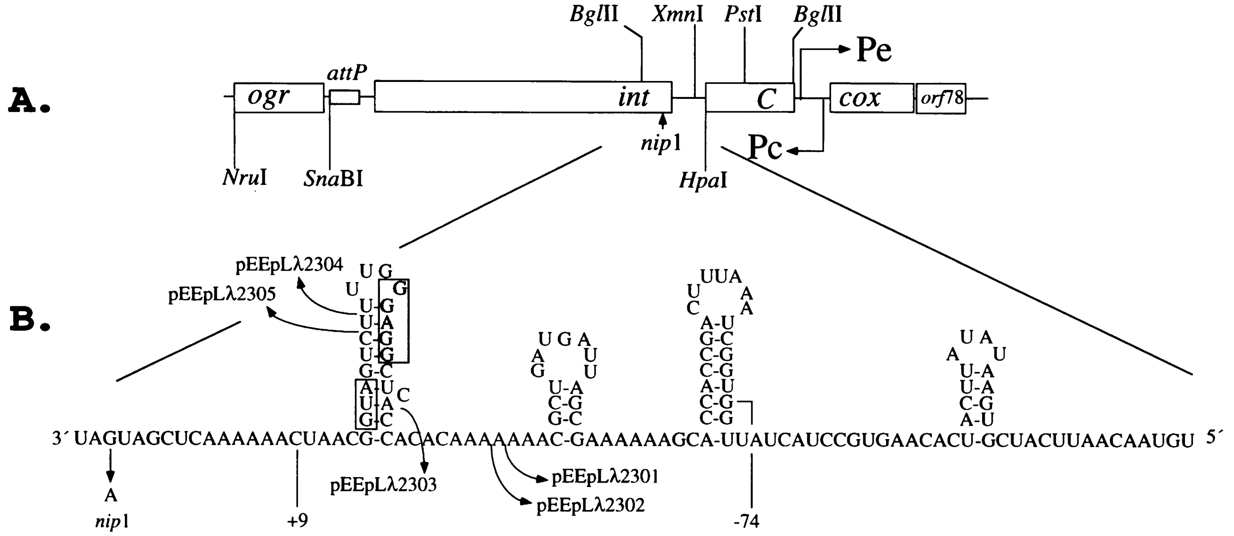

FIG. 1. (A) Schematic depiction of the genetic organization of the P2 *att*P region, indicating relevant restriction sites, the P_e and P_c promoters, and the position of the *nipl* mutation. (B) The primary structure of the transcript across the *C-int* spacer region, showing possible stem-loop structures. The shaded box indicates the bases comprising the ribosome binding site for *int*; the unshaded box indicates the initiation codon. Note that the sequence is drawn in the 3′ to 5′ direction. The curved arrows indicate the 5′ ends of the P2 sequences in the deletion plasmids generated by exonuclease III (see Materials and Methods).

zymes were obtained from Promega and used as recommended by the manufacturer.

(i) Construction of pEE2090 and pEE2091. The *Xmn*I-*Bgl*II fragment (Fig. 1) containing most of the *int-C* spacer region and the first 67 nucleotides of the *int* gene was inserted in the unique *Sma*I site of pKG1901 (24). The two orientations of the insert relative to P*gal* and *galK* are depicted in Fig. 2.

(ii) Construction of pEE2100, pEE2101, and pEE2102 (Fig. 3A). The promoterless *cat* gene (Pharmacia) was inserted in the unique *Hin*dIII site of pMG524 to yield pEE2100. The *Hpa*I-*Sna*BI fragment of P2 containing the *C-int* spacer region, the structural *int* gene, and the *att*P core sequence, was inserted in the *Sma*I site of pEE2100 to yield pEE2101. The *Hpa*I-*Nru*I fragment of P2, which contained additional downstream sequences relative to *int*, was inserted in pEE2100 in the *Sma*I site to yield pEE2102.

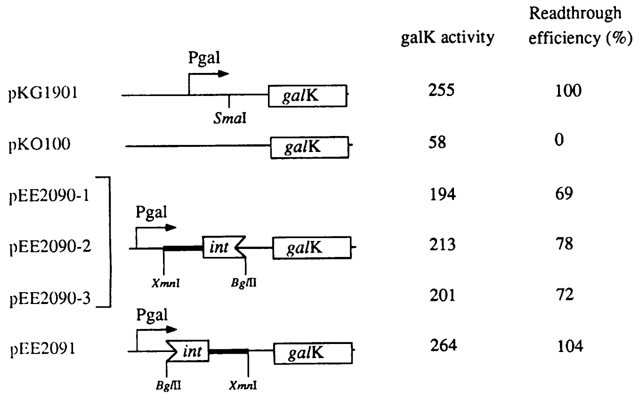

FIG. 2. Partial termination of transcription in the *int-C* spacer region. The construction of the vectors, the hosts used, and the calculation of *galK* activity units is described in Materials and Methods. Readthrough efficiency in vivo was calculated from *galK* activities by setting the activity of pKG1901 as 100% and that of pKO100 as 0%.

(iii) Construction of *int-lacZ* fusion plasmids (Fig. 4). The *Hpa*I-*Bgl*II fragment of wild-type (wt) P2 or P2 *nip1*, generated by the PCR technique, was inserted into the unique *Sma*I site of pMC1871 (32) to yield pEE2306(wt) and pEE2307(*nip1*). Sequencing verified that the coding *int* gene in each case was in-frame with that of *lacZ*.

pEE2308 was created by insertion of the *Sal*I restriction fragment of pMC1871 containing the *lacZ* gene into the *Bam*HI site of pET3 (33). pEE2309(wt) and pEE2310(*nip1*) were constructed by inserting the *Pst*I restriction fragment of pEE2306(wt) or pEE2307(*nip1*) containing the respective *int-lacZ* fusion genes into the *Bam*HI site of pET3.

pEE2501 was constructed by first inserting a restriction fragment containing the λpR promoter and the temperature-sensitive repressor gene *cI*857 in the *Pst*I site of pACYC177 (6). The *Eco*RI-*Hin*dIII fragment of pEE2101 (Fig. 3A) containing *int* was subsequently placed downstream of λpR, so that the expression of *int* was directed by this promoter.

Determination of *cat* activity. The chloramphenicol acetyltransferase (CAT) assay was performed as described in reference 12. Protein concentrations were determined by the method of Bradford, with bovine serum albumin (BSA) as a standard (5). After autoradiography, we quantified the CAT activity by cutting out the radioactive spots from the silica gel thin-layer chromatography plates and counting the radioactivity in a liquid scintillation counter.

Determination of *galK* activity. These assays were carried out as previously described (20), with the following modifications: cells containing plasmid were grown to log phase in M9 medium (27) supplemented with casamino acids and fructose as a carbon source. After lysis and reaction as described, 50-μl aliquots were placed on DE81 filters (Whatman) and washed twice with 10 ml of water each time. After completion of the assay, 50-μl aliquots from two randomly selected sample tubes were transferred to two additional filter discs. Galactokinase units are expressed as nanomoles of galactose phosphorylated per minute per milliliter of cells with an optical density at 650 nm of 1.0.

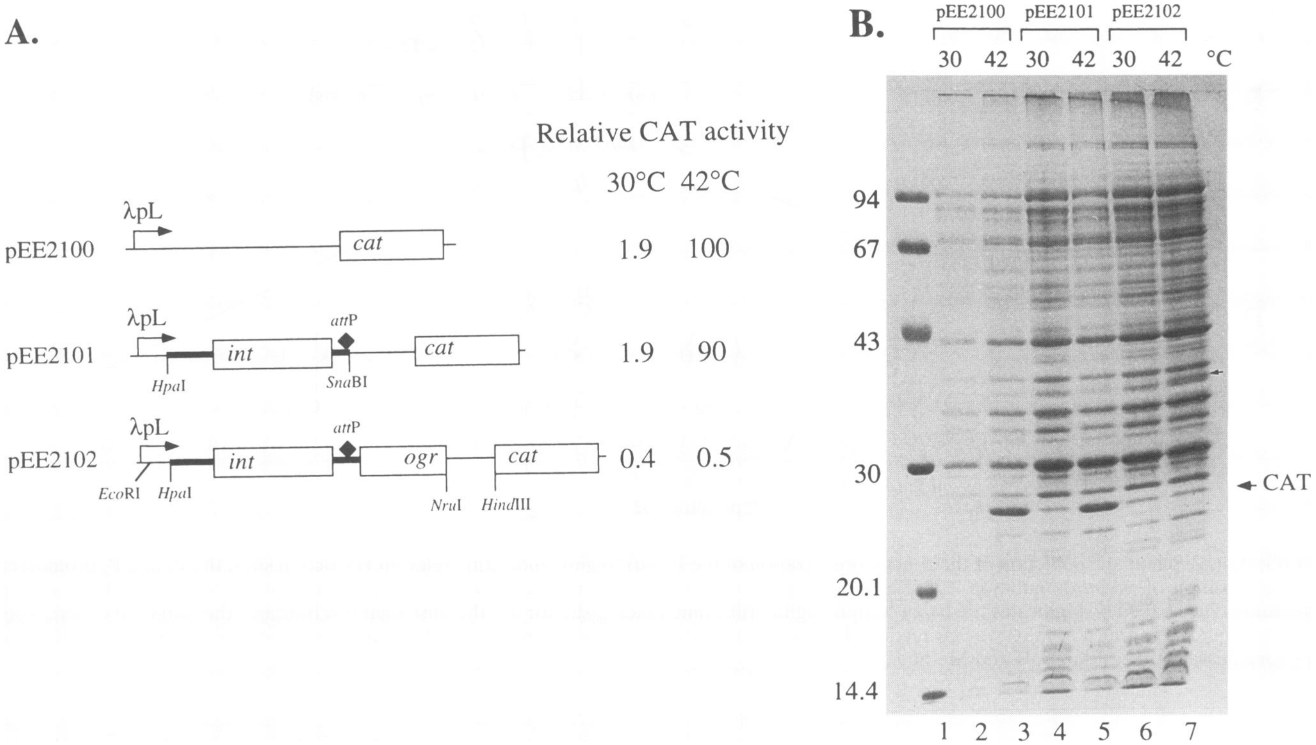

FIG. 3. (A) Constructions of the CAT reporter plasmids pEE2101 and pEE2102 and the control plasmid pEE2100 (see Materials and Methods). The relative CAT activities of cells harboring the respective plasmids are listed to the right of the drawing. (B) SDS-polyacrylamide gel stained with Coomassie blue, showing protein expression by the CAT reporter plasmids at 30 and 42°C, as indicated above lanes 2 to 7. Lane 1, molecular mass markers (phosphorylase *b*, 94 kDa; BSA, 67 kDa; ovalbumin, 43 kDa; carbonic anhydrase, 30 kDa; soybean trypsin inhibitor, 20.1 kDa; and α-lactalbumin, 14.4 kDa).

Determination of *lacZ* activity. The β-galactosidase assay was described by Miller (21) using ONPG (*o*-nitrophenyl-β-D-galactopyranoside) as a substrate.

RNA binding assay. For the RNA-binding protein blot assay, a radiolabelled RNA probe containing the ribosome binding site (RBS) of *int* was synthesized as follows: plasmid pEE2309(wt) (Fig. 4) was cleaved with *Hpa*I and used as a template for in vitro RNA synthesis with T7 RNA polymerase (Promega), in the presence of [α-^{32}P]CTP (Amersham). As a control, a radiolabelled RNA probe was synthesized in the same manner, by using *Hpa*I-cleaved pEE2308 (Fig. 4) as a template.

The proteins were resolved on duplicate 12% sodium dodecyl sulfate (SDS)-polyacrylamide gels and were visualized on one gel by staining with Coomassie brilliant blue. The other gel was equilibrated for 15 min with cold transfer buffer (25 mM Tris, 192 mM glycine, 30% methanol), after which the proteins were transferred to nitrocellulose filters. The protein blot was treated for 1 h at room temperature in binding buffer (10 mM Tris-HCl [pH 7.4], 50 mM NaCl, 1 mM EDTA, and 1× Denhardt's solution [0.02% Ficoll, 0.02% polyvinylpyrolidon, 0.03% BSA]) and was then probed in binding buffer at 40°C for 1 to 2 h with the labelled RNA. After two washes for 15 min each with binding buffer at 40°C, the blot was sealed in plastic and exposed to X-ray film for autoradiography.

Nuclease digestion analysis. Partial digestions with RNase T$_1$ were performed in 60 mM NH$_4$Cl, 10 mM magnesium acetate, 10 mM Tris-HCl [pH 7.5], and 6 mM β-mercaptoethanol. Reactions in a total volume of 15 μl containing about 1 μg of synthetic *int* mRNA (see above) were equilibrated at 30°C for 15 min with Int protein or an equal volume of Int storage buffer (0.3 M potassium phosphate [pH 7.5]). The reaction mixtures were heated to 37°C, and 0.2 U of RNase T$_1$ was added. Digestion was allowed to proceed for 2 min, and the reactions were stopped by the addition of 15 μl of 0.4 M sodium acetate [pH 5.2], 20 mM EDTA, and 1.0 μg of tRNA per μl. Following phenol treatment, ether extraction, and ethanol precipitation, the digested RNAs were analyzed by primer extension, using the primer 5′ GCG ACA GCT TCG CTT TTC TT 3′ (Innovagen, Lund, Sweden).

RESULTS

A weak transcription terminator is located between the *C* and *int* genes. In order to test if the transcript initiated by P$_c$ was partially terminated before *int*, a system composed of the reporter gene *galK* and its promoter P*gal* was employed. Insertion of this DNA fragment in the correct orientation between *galK* and its promoter resulted in a reduction of about 30% in the readthrough efficiency compared with that of the control plasmid with no insert (Fig. 2). Although this was not a drastic reduction in *galK* expression, the results were reproducible in several independent clones (pEE2090-1 to -3) and with repeated measurements. As it is known that insertion of almost any fragment will result in some polar effect on *galK* expression, we tested the effect of inserting the same DNA fragment in the opposite orientation and found that this did not reduce *galK* expression. These data suggest that, in vivo, there is a partial block in RNA synthesis through the region upstream of *int*.

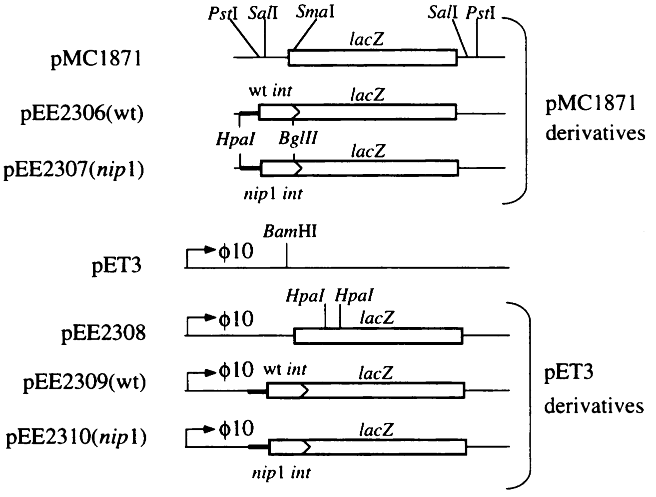

FIG. 4. Construction of *int-lacZ* fusion plasmids used to quantify *int* expression (see Materials and Methods for a detailed description of the cloning strategies).

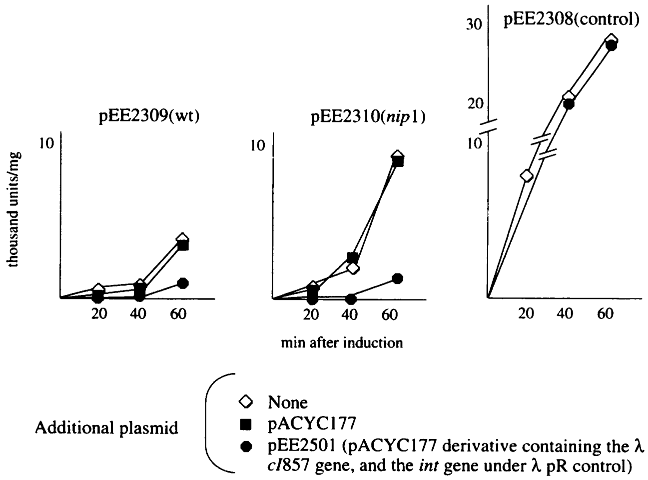

FIG. 5. Lac activities of *E. coli* JM109(DE3) harboring pEE2309(wt), pEE2310(*nipl*), or pEE2308 (control) upon induction of the T7 ϕ10 promoter by IPTG (isopropyl-β-D-thiogalactopyranoside). The individual curves in each graph depict the Lac activity in strains with the additional plasmids as indicated.

The *int-C* spacer region effectively represses *int* expression. When the *Hpa*I-*Sna*BI fragment was inserted into a λP_L expression vector (pEE2101), no Int production could be detected in SDS-polyacrylamide gels stained with Coomassie blue (Fig. 3A and B, lanes 4 and 5). The sequential removal of nucleotides between the *Hpa*I site and the *int* RBS by exonuclease III (Fig. 1) did not increase *int* expression. It has been shown that efficient expression of *int* could be achieved by removal of the RBS and initiation codon of *int* and supplying a new RBS and ATG codon, in the pET system (33, 37). The expression of *int* was further enhanced by the removal of nucleotides 4 to 18 of the *int* reading frame. In light of the finding that the transcription termination signal in this region is only partial, it seemed likely that the observed repression of *int* expression was on a posttranscriptional level, for instance in the initiation of translation, or in mRNA stability.

Evidence of *int* repression on a posttranscriptional level. In order to test this hypothesis, a reporter gene (*cat*) was inserted downstream of the *Hpa*I-*Sna*BI insert (Fig. 3A), to yield pEE2101. Cells harboring pEE2101 showed a slightly reduced level of CAT activity upon induction of λP_L compared with those harboring the control plasmid pEE2100; this was in agreement with the previous finding that transcription over the *C-int* spacer regions was partially terminated. However, on an SDS-polyacrylamide gel stained with Coomassie blue, only the *cat* gene product could be detected (Fig. 3B). Because the *cat* gene was placed downstream of *int* relative to λP_L, it was evident that the expression of *int* was being regulated at a posttranscriptional level.

When a larger P2 fragment (*Hpa*I-*Nru*I, Fig. 3A) was inserted between *cat* and λP_L, CAT activity could no longer be induced by this promoter (Fig. 3A and B, lanes 6 and 7). This is in agreement with the location of the natural transcriptional terminator of *int* just downstream of the *Sna*BI site (36). However, the interesting observation was made that in the crude extracts of cells harboring pEE2102 and induced at 42°C, a new band of 38 kDa can be detected, located just above a band present under uninduced conditions, which thus could be the Int protein (Fig. 3B, lane 7). Although we have not further investigated this, the data suggest that inclusion of the natural termination signal of the *int* transcript may have a stabilizing effect on mRNA, resulting in a prolonged half-life and increased protein levels.

Autoregulation of *int*. In order to explore the possibility that the Int protein acts as a translational repressor of its own gene, we have cloned the *Hpa*I-*Bgl*II fragment (Fig. 1) of P2 wt and *nipl* into plasmid pMC1871, which contains a promoterless *lacZ* gene, giving rise to an *int-lacZ* fusion preceded by the *C-int* spacer region (Materials and Methods). The product of the fused gene should be a protein containing 67 amino acid residues of the N-terminal end of the Int protein coupled to residue 8 of LacZ via a linker of two residues.

Neither plasmid pEE2306(wt), containing the wt *int-lacZ* fusion, or plasmid pEE2307(*nipl*), containing the *nipl-lacZ* fusion, expressed any detectable *lacZ* activity. Thus, under the conditions used, we find no evidence for an *int*-specific promoter in the *C-int* spacer region and no support for the hypothesis that the *nipl* mutation should generate a promoter function. We have not found any promoter-like activity within the *C* gene itself, which would allow *int* expression in the absence of *C* expression (data not shown). However, these experiments do not consider the possibility of an activator-dependent promoter for *int*.

To analyze the possible effects of *nipl* on the repressive effects of the *C-int* spacer region on *int* expression, the fusion proteins with the spacer regions were put under the control of the inducible ϕ10 promoter of T7 (Materials and Methods). Sixty minutes after induction, it was found that cells harboring the plasmid pEE2310(*nipl*) showed a fivefold higher *lacZ* activity than those harboring pEE2309(wt) (Fig. 4 and 5). This suggests that the *nipl* mutation affects either the translational initiation of the *int* transcript and/or the readthrough of the latter over the spacer region, because the same promoter was present in both constructs. The repression of *int* expression by the *int-C* spacer region is reflected in the low levels of *lacZ* activity yielded by both pEE2309(wt) and pEE2310(*nipl*), compared with a control plasmid which does not contain this region (pEE2308).

To test the effect of Int in *trans* on this system, a compatible plasmid expressing the *int* gene (pEE2501; see Materials and Methods) was used. A drastic decrease in *lacZ* expression upon

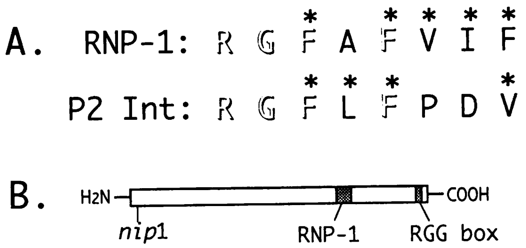

FIG. 6. (A) A comparison of the RNP-1 motif with a potential RNA-binding site of P2 Int. The letters refer to the single-letter codes for amino acids. The conserved residues are outlined; the asterisks indicate aromatic residues which contribute to the hydrophobic core. (B) Schematic representation of structural motifs identified in the P2 Int protein. The open bar indicates the coding region. Putative functional domains, the RNP-1 site, and the RGG tripeptide are indicated.

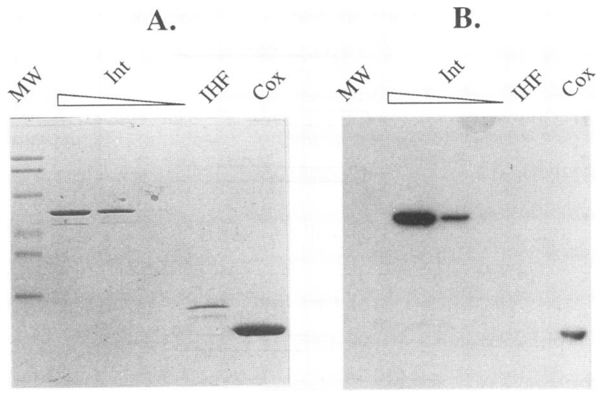

FIG. 7. RNA binds to immobilized Int protein. (A) SDS-polyacrylamide gel stained with Coomassie blue showing 1 μg of each prestained protein size marker (MW) (BioRad) (phosphorylase *b*, 110 kDa; BSA 84 kDa; ovalbumin, 47 kDa; carbonic anhydrase, 33 kDa; trypsin inhibitor, 24 kDa; lysozyme 16 kDa), purified Int (2 μg, 1 μg, and 0.5 μg), 1 μg of purified IHF, and 4 μg of purified Cox. (B) Analysis of a protein blot assay. A duplicate SDS-polyacrylamide gel was transferred to nitrocellulose membrane, incubated with radioactive Int-specific mRNA, washed, and autoradiographed. Preparation of the mRNA and the RNA-binding protein blot assay are described in Materials and Methods.

induction of φ10 was measured in the strain harboring both plasmids pEE2309(wt) and pEE2501 and in the strain harboring both pEE2310(*nip1*) and pEE2501. The control strain harboring pEE2308 and pACYC177 retained high inducible *lacZ* activity (Fig. 5).

These results suggest the following: (i) the Int protein negatively regulates the expression of its own gene; (ii) this autoregulation lies on a posttranscriptional level; and (iii) although the *Hpa*I-*Bgl*II region includes the sequences which are necessary for *int* autoregulation, the *nip1* mutation does not have an effect on the latter.

Int protein has potential RNA-binding motifs. In order to test the hypothesis that Int regulates the expression of its own gene on the translational level, we examined the predicted primary structure of Int and found a stretch of eight residues in good agreement with the RNP-1 conserved sequence of a family of RNA-binding proteins (15) (Fig. 6A and 6B). This rather diverse family of proteins includes small nuclear ribonucleoprotein, splicing factors, heterogenous nuclear ribonucleoprotein, translation factors, and the *E. coli* rho factor. In addition to this, several RNA-binding proteins have recently been found to contain an arginine- and glycine-rich domain that contains a cluster of the tripeptide repeat called the RGG box (16). This motif has been found in a number of nuclear and nucleolar RNA-binding proteins and has been demonstrated to have RNA-binding activity (8). The Int protein contains one copy of the RGG tripeptide, and Fig. 6B shows its position near the carboxyl end of Int.

Int protein binds to RNA in vitro. In order to test whether P2 Int is an RNA-binding protein, we used an RNA-binding protein blot assay. With this approach, the purified proteins Int, IHF, and Cox were immobilized on nitrocellulose after SDS-polyacrylamide gel electrophoresis, probed with radioactively labelled RNA, and visualized by autoradiography. For this experiment, in vitro-synthesized mRNA specified by *Hpa*I-*Bgl*II sequences including the RBS of *int* was used (see Materials and Methods). Figure 7 shows that purified Int bound RNA (and to a certain extent, even purified Cox). The lack of binding to the proteins in the molecular mass marker lane (prestained phosphorylase *b*, BSA, ovalbumin, carbonic anhydrase, trypsin inhibitor, and lysozyme) or to IHF demonstrate the specificity of the binding of Int to RNA. On the basis of this experiment, we conclude that Int is an RNA-binding protein in vitro and thus has the capacity to also bind RNA in vivo. The specificity of Int binding could not be demonstrated by this assay, as Int also bound non-*int* RNA. Furthermore, Int-*int* RNA binding could not withstand heparin challenge (data not shown). Therefore, we proceeded with an RNA footprint assay.

The Int binding site on the *int* gene mRNA. Standard footprinting techniques and primer extension were used to locate the binding site of Int protein on its own transcript. RNase T_1 digestion of the *int* gene mRNA in the presence or absence of Int protein shows that the protein can specifically protect from nuclease digestion a region near and upstream of the *int* initiation codon (Fig. 8). The region of the mRNA protected by Int from RNase digestion extends from G (−74) to C (+9). This region may extend further, as the closest unprotected bases outside the protected region are U (−77) and A (+15). The protection of the transcript by Int protein was quantified by densitometric scanning of the gel shown in Fig. 8. The extent of protection was found to depend on the amount of Int present, and the region from G (−74) to C (+9) is strongly protected (50 to 80%) by 5 μg of Int in a reaction volume of 15 μl.

DISCUSSION

Int is encoded in the P2 genome adjacent to the attachment site, reflecting the common pattern of recombinase genes close to their sites of action. The functions of the recombinases include a wide range of biological processes, such as phage integration and excision, switching of gene expression exemplified by fimbrial phase variation, plasmid copy number regulation, splicing of genes at the DNA level (immunoglobulin and T-cell receptor genes), resolution of transposition intermediates, and monomerization of multimeric bacterial plasmids and chromosomes (4, 10, 25, 29). It is usually essential for their functions that site-specific recombination reactions are well regulated and precise, since overproduction of the recombinases often leads to enhanced recombination functions which would be harmful to the cell. In contrast to this, it has been shown that integrative recombination in vitro in P2 is inhibited by high Int concentrations (38). Therefore, it is not surprising to find sophisticated strategies for the regulation of

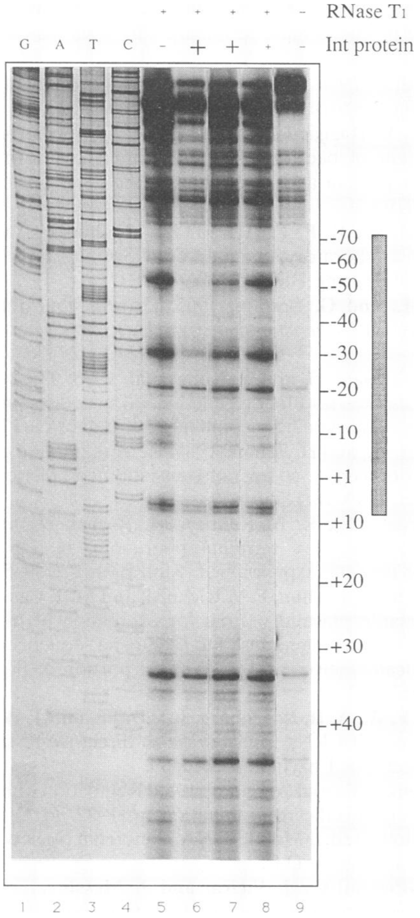

FIG. 8. Footprinting of Int protein bound to *int* gene mRNA. Partial RNase T_1 cleavage of *int* RNA was carried out in the absence (lane 5) or presence (lanes 6, 7, and 8) of decreasing Int protein concentrations (5 μg, 2.5 μg, or 1.25 μg per reaction mixture), and the digested RNA was analyzed by primer extension.

expression of recombination genes. In the case of λ *int*, this involves transcription from two different promoters, the cII-dependent P_{int} and the distant P_L (34), as well as retroregulation (1, 13, 22, 30, 31).

The mode by which *int* regulation occurs in P2 is unknown. Although the *C* and *int* genes are cotranscribed from the P_c promoter, in vitro recombination assays have indicated that an uncontrolled overproduction of Int is deleterious to the integrative process (38). We have shown that the extent of *int* expression is controlled at three different levels. First, a partial transcription terminator situated in the spacer region between the *C* and *int* genes reduces transcriptional readthrough. Second, the initiation of translation of *int* is blocked, possibly because of the location of the RBS and ATG codon of the gene in a putative stem-loop structure. Intramolecular base pairing that involves the ribosome binding site and the initiation codon are known to slow down or even inhibit translational initiation (11, 19). Interestingly, the *nip1* mutation was found to alleviate this blockage, suggesting that this mutation (a G to A substitution at the 22nd coding nucleotide of *int*) results in an alternatively looped mRNA structure which could increase the accessibility of the RBS to ribosomes.

Our data indicate that a third mechanism and probably the major mode of controlling *int* gene expression in P2 is that of autoregulation. When Int protein is provided in *trans* to cells harboring a plasmid containing an *int-lacZ* fusion gene under the direction of the ϕ10 promoter of the phage T7, *lac* activity upon induction of the promoter by IPTG is effectively reduced to basal levels. Although the experiments depicted in Fig. 5 show that the *nip1* mutation is amenable to autoregulation by wt Int, the situation in vivo is that Nip1 protein and not wt Int is produced. Therefore, the possibility remains that Nip1 protein is less effective in regulating the *nip1* gene.

Several lines of evidence presented here point strongly to regulation at a posttranscriptional level. To further test this hypothesis, we performed RNA-binding protein blot assays which showed that the P2 Int protein actually binds to RNA. The site of Int protein binding to its own mRNA was located to a region extending from −74 to +9. This region does not include any arm-type or core-type Int recognition sequences (37), suggesting that the mode of Int recognition of, or binding to, single-stranded RNA differs from its recognition of, or binding to, DNA. Negative autoregulation at the translational level was first demonstrated for gene 43 of bacteriophage T4, and since then several other systems have been described (11, 19).

The direct conclusion from the RNA binding assays and footprinting experiments is that the binding of Int protein to its own transcript may block ribosome binding and thereby initiation of translation. The observation that exonuclease III deletion of sequences within this region (Fig. 1) did not significantly increase *int* expression suggests that the bases −16 to +9 are sufficient for Int recognition of the mRNA. Our results suggest two different mechanisms by which the initiation of translation of the *int* gene may be controlled, namely by occlusion of the RBS to ribosomes by Int protein, or by secondary mRNA structure. A more intriguing implication of the RNA-binding properties of P2 Int is that, analogously to the spliceosomes of eukaryotic systems, the integrative/excisive intasome of a bacteriophage has RNA as an essential component, either structurally or catalytically. This theory, which is supported by the observation that another component of the intasome, Cox, also binds RNA, remains to be resolved by direct experimentation.

ACKNOWLEDGMENTS

This work was supported by grant 72 from the Swedish Medical Research Council (E.H.-L.) and by the memorial funds of Selma and Fritz Ringström, Thyr and Ture Stenemark, Ruth Trossbeck, and Lars Hierta (A.Y.).

REFERENCES

1. **Belfort, M.** 1980. The *c*II-independent expression of the phage λ *int* gene in RNaseIII-defective *E. coli*. Gene **11**:149–155.
2. **Bertani, G.** 1962. Multiple lysogeny from single infection. Virology **13**:131–139.
3. **Bertani, L. E.** 1968. Abortive induction of bacteriophage P2. Virology **36**:87–103.
4. **Blakely, G., S. Colloms, G. May, M. Burke, and D. Sherratt.** 1991. *Escherichia coli XerC* recombinase is required for chromosomal segregation at cell division. New Biol. **3**:789–798.
5. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248–254.
6. **Chang, A. C. Y., and S. N. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. **134:** 1141–1156.
7. **Dodd, I. B., B. Kalionis, and J. B. Egan.** 1990. Control of gene expression in the temperate coliphage 186. VIII. Control of lysis and lysogeny by a transcriptional switch involving face to face promoters. J. Mol. Biol. **214**:27–37.

8. **Dreyfuss, G., M. J. Matunis, S. Pinol-Roma, and C. G. Burd.** 1993. hnRNP proteins and the biogenesis of mRNA. Annu. Rev. Biochem. **62:**289–321.
9. **Echols, H., and G. Guarneros.** 1983. Control of integration and excision, p. 75–92. *In* R. W. Hendrix, J. W. Roberts, F. W. Stahl, and R. A. Weisberg (ed.), Lambda II. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
10. **Gellert, M., and H. Nash.** 1987. Communication between segments of DNA during site-specific recombination. Nature (London) **325:**401–404.
11. **Gold, L.** 1988. Posttranscriptional regulatory mechanisms in *Escherichia coli*. Annu. Rev. Biochem. **57:**199–233.
12. **Gorman, C. M., L. F. Moffat, and B. M. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2:**1044–1051.
13. **Guarneros, G., C. Montanez, T. Hernandez, and D. Court.** 1982. Posttranscriptional control of bacteriophage λ *int* gene expression from a site distal to the gene. Proc. Natl. Acad. Sci. USA **79:**238–242.
14. **Heaphy, S., M. Singh, and M. J. Gait.** 1978. Effect of single amino acid changes in the region of the adenylation site of T4 RNA ligase. Biochemistry **26:**1688–1696.
15. **Kenan, D. J., C. C. Query, and J. D. Keene.** 1991. RNA recognition: towards identifying determinant of specificity. Trends Biochem. Sci. **16:**214–220.
16. **Kiledjian, M., and G. Dreyfuss.** 1992. Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box. EMBO J. **11:**2655–2664.
17. **Landy, A.** 1989. Dynamic, structural, and regulatory aspects of λ site-specific recombination. Annu. Rev. Biochem. **58:**913–949.
18. **Ljungquist, E., and L. E. Bertani.** 1983. Properties and products of the cloned *int* gene of bacteriophage P2. Mol. Gen. Genet. **192:**87–94.
19. **McCarthy, J. E., and C. Gualerzi.** 1990. Translational control of prokaryotic gene expression. Trends Genet. **6:**78–85.
20. **McKenney, K., H. Shimatake, D. Court, U. Schmeissner, C. Brady, and M. Rosenberg.** 1981. A system to study promoter and terminator signals recognized by *Escherichia coli* RNA polymerase, p. 383–415. *In* J. S. Chirikjian and T. S. Papas (ed.), Gene amplification and analysis. Elsevier North Holland, New York.
21. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
22. **Montanez, C., J. Buero, U. Schmeissner, D. Court, and G. Guarneros.** 1986. Mutations of bacteriophage lambda that define independent but overlapping RNA processing and transcription termination sites. J. Mol. Biol. **191:**29–37.
23. **Norrander, J., T. Kempe, and J. Messing.** 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene **26:**101–106.
24. **Rosenberg, M., A. B. Chepelinsky, and K. McKenney.** 1983. Studying promoters and terminators by gene fusion. Science **222:**734–739.
25. **Sadowski, P.** 1986. Site-specific recombinases: changing partners and doing the twist. J. Bacteriol. **165:**341–347.
26. **Saha, S., E. Haggård-Ljungquist, and K. Nordström.** 1987. The *cox* protein of bacteriophage P2 inhibits the formation of the repressor protein and autoregulates the early operon. EMBO J. **6:**3191–3199.
27. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
28. **Sasaki, I., and G. Bertani.** 1965. Growth abnormalities in Hfr derivatives of *Escherichia coli* strain C. J. Gen. Microbiol. **50:**365–376.
29. **Sato, T., Y. Samori, and Y. Kobayashi.** 1990. The *cisA* cistron of *Bacillus subtilis* sporulation gene *spoIVC* encodes a protein homologous to a site-specific recombinase. J. Bacteriol. **172:**1092–1098.
30. **Schindler, D., and H. Echols.** 1981. Retroregulation of the *int* gene of bacteriophage λ: control of translation completion. Proc. Natl. Acad. Sci. USA **78:**4475–4479.
31. **Schmeissner, U., K. McKenney, M. Rosenberg, and D. Court.** 1984. Removal of a terminator structure by RNA processing regulates *int* gene expression. J. Mol. Biol. **176:**39–53.
32. **Shapira, S. K., J. Chou, F. V. Richard, and M. J. Casadaban.** 1983. New versatile plasmid vectors for expression of hybrid proteins coded by a cloned gene fused to *lac*Z gene sequences encoding an enzymatically active carboxy-terminal portion of β-galactosidase. Gene **25:**71–82.
33. **Studier, F. W., A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff.** 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. **185:**60–89.
34. **Thompson, J. F., and A. Landy.** 1989. Regulation of bacteriophage lambda site-specific recombination, p. 1–22. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
35. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.
36. **Yu, A., L. E. Bertani, and E. Haggård-Ljungquist.** 1989. Control of prophage integration and excision in bacteriophage P2: nucleotide sequences of *int* gene and *att* sites. Gene **80:**1–12.
37. **Yu, A., and E. Haggård-Ljungquist.** 1993. Characterization of the binding sites of two proteins involved in the bacteriophage P2 site-specific recombination system. J. Bacteriol. **175:**1239–1249.
38. **Yu, A., and E. Haggård-Ljungquist.** 1993. The Cox protein is a modulator of directionality in bacteriophage P2 site-specific recombination. J. Bacteriol. **175:**7848–7855.