Cloning, Sequencing, and Genetic Characterization of Regulatory Genes, \textit{rina} and \textit{rinB}, Required for the Activation of Staphylococcal Phage \textit{phi}11 \textit{int} Expression

ZH-I-HAI YE AND CHIA Y. LEE*

Department of Microbiology, Molecular Genetics and Immunology, University of Kansas Medical Center, Kansas City, Kansas 66103

Received 2 November 1992/Accepted 11 December 1992

The \textit{int} gene of staphylococcal bacteriophage \textit{phi}11 is the only viral gene responsible for the integrative recombination of \textit{phi}11. To study the regulation of \textit{int} gene expression, we determined the 3' end of the transcript by S1 mapping. The presumed promoter is located just 22 nucleotides upstream of the \textit{int} open reading frame in a region which is conserved between \textit{phi}11 and a closely related staphylococcal phage, L54a. To clone the possible regulatory gene, a vector which contained the reporter gene, \textit{xylE}, of \textit{Pseudomonas putida} under the control of the \textit{phi}11 \textit{int} promoter was constructed. Subsequently, a 2-kb DNA fragment from the \textit{phi}11 genome, which mapped distal to the \textit{int} gene, was shown to increase the \textit{xylE} activity from the \textit{int} promoter. Sequencing and subsequent deletion analysis of the 2-kb fragment revealed that two \textit{phi}11 regulatory genes, \textit{rina} and \textit{rinB}, were both required to activate expression of the \textit{int} gene. Northern (RNA) analysis suggested that the activation was, at least partly, at the transcriptional level. In addition, one of these regulatory genes, \textit{rina}, was capable of activating L54a \textit{int} gene transcription.

Staphylococcal bacteriophages \textit{phi}11 and L54a are serogroup B, lytic type III phages originally isolated from \textit{Staphylococcus aureus} 8325 (20) and PS54 (5), respectively. They are temperate phages capable of both lytic and lysogenic growth. In the lysogenic life cycle, the phage DNA recombines with the host chromosome by a mechanism of site-specific recombination occurring between the phage attachment site, \textit{attP}, and the bacterial attachment site, \textit{attB} (16, 17). The \textit{attB} sites of both phages have been mapped on the genetic map of \textit{S. aureus} NTCC8325. The \textit{phi}11 \textit{attB} site is located between the \textit{purB} and \textit{metA} loci, whereas the L54a \textit{attB} site is mapped within the lipase gene (\textit{geh}) near \textit{tet} (22, 23).

Our previous studies on site-specific recombination of the two staphylococcal phages have demonstrated that two phage-encoded proteins, integrase (\textit{Int}) and excisionase (\textit{Xis}), are involved in the recombination reactions (30, 31). In integrative recombination, only \textit{Int} is required, while both \textit{Int} and \textit{Xis} are required in excisive recombination. This feature is similar to that of coliphage \textit{\lambda} (28), indicating that \textit{phi}11 and L54a may control their site-specific recombination by differential expression of \textit{int} and \textit{xis} genes. In the \textit{\lambda} system, the expression of \textit{int} is controlled by complex mechanisms which include activation by \textit{ci}, antitermination by N protein, repression by \textit{cI}, and retroregulation at the \textit{sb} site (see references 6 and 28 for reviews).

The DNA regions involved in the site-specific recombination of \textit{phi}11 and L54a have been cloned and sequenced (30, 31). The two regions are similar in genetic organization, with the gene order \textit{attP-int-xis}. The \textit{int} gene is transcribed toward the \textit{attP} site, whereas the \textit{xis} gene is transcribed away from the \textit{attP} site. Sequence comparison between the two phages, however, showed very little homology in the \textit{int} and \textit{xis} coding regions. Surprisingly, a high degree of homology was found at a region between the translational start

\begin{itemize}
  \item \textit{Corresponding author.}
\end{itemize}
were purchased from New England BioLabs, Inc. (Beverly, Mass.), and Bethesda Research Laboratories, Inc. (Gaithersburg, Md.). Restriction enzyme digestions were performed under conditions recommended by the supplier. The large fragment of DNA polymerase (Klenow fragment) and a DNA sequencing kit including dideoxy nucleoside triphosphates were purchased from U.S. Biochemical Corp. (Cleveland, Ohio). [α-32P]dCTP was purchased from New England Nuclear Corp. (Boston, Mass.).

**Recombinant DNA techniques.** General DNA manipulations were performed as described by Sambrook et al. (26). Deletion mutagenesis in vitro was performed by the exonuclease III deletion procedure (7, 10). Rapid small-scale DNA purification was done by the method of Holmes and Quigley (11).

**DNA sequence analysis.** The 2.0-kb insert containing the *rint* genes was cloned into bacteriophage M13 derivatives mp18 and mp19 (29). The unidirectional exonuclease III deletion procedure was used to generate overlapping deletions of the DNA fragment for sequencing as described previously (30). DNA sequencing was carried out by the dyeoxy-chain termination method of Sanger et al. (27).

**S1 mapping.** Total cellular RNA was isolated from strains containing pCL2112, pYL306, or pYL240 as described previously (26). The transcription start site of the *int* gene was determined by S1 mapping (1). A 32P-labeled single-stranded DNA probe was made by using recombinant bacteriophage M13 mp18 which contained the 0.5-kb *BglII-AccI* fragment from pCL2112 (see Fig. 2) at the *BamHI-AccI* site. The 0.5-kb DNA probe spanning the upstream region of the *int* gene was generated with [α-32P]dATP by using the Klenow fragment of DNA polymerase I, which yielded a specific activity of 7.0 × 107 cpm/μg. The 32P-labeled probe (106 cpm) was hybridized to 150 μg of RNA at 30°C for 16 h and then subjected to digestion with 330 U of S1 nuclease at 42°C for 2 h. The protected DNA fragment was ethanol precipitated, analyzed by electrophoresis in a 6% polyacrylamide-8 M urea gel, and visualized by autoradiography.

**Northern (RNA) hybridization.** Total cellular RNA was isolated from cells containing pLC4, pYL306, or pYL240 and analyzed by Northern blotting (26). RNA samples were fractionated on a 1% agarose gel containing formaldehyde and transferred to a nitrocellulose filter. Blots were prehybridized for 2 h at 42°C in a solution containing 50% formamide, 5× SSPE, 5′ Denhardt’s solution, 0.1% sodium dodecyl sulfate (SDS), and 100 μg of denatured salmon sperm DNA per ml. Hybridization was carried out at 42°C for 20 h in the same solution with 106 cpm of single-stranded DNA probe specific for the *int* gene (a 530-bp *AccI-HindIII* DNA fragment containing the N-terminal half of the *int* gene [30]) made with [α-32P]dATP by using the Klenow fragment of DNA polymerase I. Blots were washed once in 2× SSPE–0.1% SDS at room temperature for 20 min and then three times in 0.2× SSPE–0.1% SDS at 65°C for 20 min. Blots were exposed to X-ray film with an intensifying screen at −70°C.

**Construction of transcriptional fusions.** The *xylE* gene from *Pseudomonas putida* has been successfully used as a reporter gene in *Bacillus subtilis* (25) and in *S. aureus* (21). To characterize the genetic regulation of the *int* gene expressed from both L54a and φ11, we constructed two transcriptional fusion plasmids by fusing each *int* promoter upstream of the promoterless *xylE* gene on plasmid pLC4. Plasmid pYL304, containing an L54a *Pint* (*int* promoter)-*xylE* fusion, was constructed by inserting the 1.0-kb *BamHI-HindIII* fragment of pCL1813 (31) containing the L54a *int* gene promoter into the *BamHI* and *HindIII* sites of plLC4. Plasmid pYL306, a φ11 *Pint-xylE* fusion plasmid, was constructed by inserting the 0.6-kb Sau3A fragment of plasmid pCL2112 (30) containing the φ11 *int* gene promoter into the *BamHI* site of plLC4. The orientation of the promoter was ensured by restriction analysis. The partial restriction map of pYL306 is shown in Fig. 1. To construct a φ11 library, the multiple cloning site from pUC18 was inserted into the *NruI* site of pYL306 by blunt-end ligation. Sau3A partially digested DNA fragments from φ11 (4 to 6 kb) were ligated to the *BamHI* site of the multiple cloning site. The catechol 2,3-dioxigenase (the gene product of *xylE*) activities of strains containing pYL306 or the recombinants were assayed as previously described (33).

**Nucleotide sequence accession number.** The GenBank accession number for the sequences reported in this paper is L07580.

**RESULTS**

**Mapping of initiation site for *int* transcription.** Previously, we reported that the DNA sequence of a region between the translational start sites of the *int* and *xis* genes was conserved between phase L54a and phase φ11 (30). This finding suggests that the conserved region may contain the potential regulatory sites for expression of the *int* or *xis* gene. Furthermore, by analogy to the *λ* system in which the *int* gene is activated by cI at the *int* gene promoter (*P*), it is conceivable that there exists an activatable promoter in φ11 similar to that of the *λ* system. Since we have shown previously that a plasmid, such as pCL2112, carrying only the *xis-int-attP* region was able to integrate into the host chromosome at a low frequency (30, 32), we postulate that the φ11 *int* gene is expressed constitutively at a low level from a proximal promoter. To determine whether there is a φ11 *int* gene promoter located in the conserved region, total cellular RNA isolated from strain RN4220 containing pCL2112 (30) was subjected to S1 mapping. The result revealed one major protected fragment (Fig. 2A). The size of the fragment, determined by comparison with a sequencing ladder, was 88 nucleotides with respect to the *AccI* site. Thus, the transcriptional initiation site was at the A residue just upstream of the ribosome-binding site of the *int ORF* (open reading frame). At 10 nucleotides upstream of the start site we found
VOL. 175, 1993

The yellow treatment fragment is indicated by an arrowhead. The sequences are separated by a thin line.

FIG. 2. S1 mapping of the int transcriptional initiation site. (A) After treatment with S1 nuclease, the protected fragment was separated by electrophoresis in a 6% sequencing gel. The protected fragment is indicated by an arrowhead. The size of the protected fragment, with respect to the AccI site, was estimated by comparison with the simultaneously run G and A sequencing reactions. (B) The broad arrow indicates the extent and the orientation of the single-stranded DNA (ss-DNA) probe used in S1 mapping. The vertical arrow represents the transcriptional start site. Consensus -10 promoter and Shine-Dalgarno (S-D) sequences upstream of the Int ORF are underlined. Inverted arrows indicate the inverted repeat sequences. The partial sequence of the N terminus of integrase is also shown.

Materials and Methods

DNA library were first constructed in E. coli LE392 and then transformed into S. aureus RN4220. The transformants were screened for xylE expression by spraying with 0.5 M catechol. From about 500 colonies, 2 yellow colonies were found. The recombinant plasmids were isolated and designated pYL228 and pYL230. Strain RN4220 carrying either one of the recombinant plasmids showed a positive XylE phenotype (6.0 to 6.5 mU; Fig. 3). The DNA fragments from the recombinant plasmids were both mapped in the K-L region of the φ11 EcoRI restriction map, between the early and late regions (see Fig. 8) (14, 19).

To confirm that the cloned DNA fragment contained the regulatory gene which activated int expression and to determine whether the activation is at the transcriptional or translational level, Northern analysis was performed. Single-stranded DNA specific for the int gene as described in the Materials and Methods was used for hybridization. The result shows that the int mRNA was elevated substantially by the presence of the cloned fragment (Fig. 4), suggesting

Cloning of the φ11 regulatory gene for int expression. The fact that P1 is located at the conserved region just upstream of the gene strongly implies that the regulatory site for the int gene is also in the conserved region. Thus, this result provides us the basis for using the conserved region to study the regulation of int gene expression. A transcriptional fusion plasmid, pYL306, was constructed by ligating the conserved region to the promoterless xylE gene as described in Materials and Methods (Fig. 1). Strain RN4220(pYL306) exhibited no XylE activity (Fig. 3) and did not give rise to yellow colonies when sprayed with 0.5 M catechol (25). The fusion plasmid pYL306 was then used as a vector to construct a φ11 genomic library. Plasmids representing the φ11

FIG. 3. XylE activity of the recombinant plasmid carrying the int activation gene. XylE activity was assayed as described in Materials and Methods and was expressed in milliunits. The DNA fragments from the recombinants containing the activation genes were mapped on the EcoRI restriction map (bold letters) of the φ11 genome (19). Numbers below the restriction map indicate sizes in kilobases. E, EcoRI; H, HindIII; B, BglII; S, Sau3A.

FIG. 4. Northern blot analysis of the int fusion mRNA. (A) Total cellular RNA was run on a 1% formaldehyde–agarose gel and stained with ethidium bromide. Lanes: 1, RNA from RN4220 carrying pLC4; 2, RNA from RN4220 carrying pYL306; 3, RNA from RN4220 carrying pYL240. (B) The RNA samples from panel A were subjected to Northern blot analysis with the single-stranded DNA probe specific for the int transcript. Sizes are indicated in kilobases.
that the activation is most likely exerted at the transcriptional level. The sizes of the two bands shown in lane 3 of Fig. 4B (a 9.6-kb band and a larger one) are much larger than the estimated size for the Pint-xylE transcript. The longer transcript may be due to the lack of an efficient termination site downstream of the xylE gene.

DNA sequencing of the rin genes. The 2.0-kb DNA fragment containing the regulatory gene was recloned to bacteriophage M13 vectors mp18 and mp19 and subjected to sequence analysis. The entire fragment was sequenced on both strands as described in Materials and Methods. Four ORFs, ORF1 to ORF4, which could encode for proteins of 140, 68, 62, and 72 residues, respectively, were found. To determine which of these ORFs was required for activation of the int gene, deletions with various lengths were generated by exonuclease III deletion in pYL270. The resultant plasmids were transformed into strain RN4220 and assayed for XylE activity. The results shown in Fig. 5 indicate that both ORF1 and ORF3 were required for activation of the int gene. ORF1 and ORF3 were designated rinA and rinB (regulatory for int gene expression), respectively. The nucleotide and deduced amino acid sequences of the rinA and rinB genes are shown in Fig. 6. The rinA ORF was preceded by a typical ribosome-binding site located 6 bp upstream of its starting codon, whereas the rinB gene was not. In addition, no consensus promoter sequence was found further upstream of either gene. To test whether the cloned DNA fragment contains the rin gene promoter, the cloned DNA fragment from pYL240 was recloned in the opposite orientation on the same vector. No XylE activity was observed, indicating that there is no promoter specific for the rin genes (result not shown). Therefore, the rin gene must be driven by a promoter from the vector.

Computer analysis of the deduced amino acid sequence showed that RinA protein was a basic protein with a pI of 9.96, suggesting that it might be a DNA-binding protein, and that the RinB protein is an acidic protein with a pI of 4.9. Neither protein was shown to have significant homology to any known regulatory protein.

Rin activates P1. To examine whether Rin activates int expression at P1 or at other promoter upstream, we performed S1 mapping by using RN4220 harboring either pYL240 or pYL306. In both cases, the 5' ends of the transcripts were all mapped to the same nucleotide as that of pCL2112, indicating that the int gene was transcribed from the P1 when activated by Rin (results not shown). To further confirm this conclusion, we deleted the Pint from the Pint-xylE fusion of pYL272 (Fig. 5). The result (not shown) showed that there was no detectable XylE activity in strain RN4220 harboring the deleted plasmid, indicating that there are no promoters other than P1 that can be activated by Rin.

φ11 rinA cross-activates L54a int. Since a highly conserved
was able to activate the L54a int gene. However, only one of the plasmids were assayed for activity. 

To determine the region responsible for activating the int gene, we cloned and sequenced the L54a DNA fragment containing the int gene. The fragment was identified by screening a λ library for activation of the int promoter fused to a reporter gene, xylE. The activation was confirmed by Northern blot analysis, which also showed that the activation was most likely exerted at the translational level. The result from the S1 nuclease protection assay shows that transcription of the int gene starts 22 nucleotides upstream of the translational start site. Although a -10 consensus sequence was found in the P1 promoter, no -35 consensus sequence was found. This feature is common to the genes which require an additional activator for effective activation.

**DISCUSSION**

We cloned and sequenced the L54a DNA fragment containing the regulatory regions required for activation of the int gene. The int genes were identified by screening a λ library for activation of the int promoter fused to a reporter gene, xylE. The activation was confirmed by Northern blot analysis, which also showed that the activation was most likely exerted at the translational level. The result from the S1 nuclease protection assay shows that transcription of the int gene starts 22 nucleotides upstream of the translational start site. Although a -10 consensus sequence was found in the P1 promoter, no -35 consensus sequence was found. This feature is common to the genes which require an additional activator for effective activation.
transcription. For example, the λ CII-dependent promoters P₁ and P₄E lack a good −35 sequence. Activation by CII requires binding of CII to a direct repeat within the −35 region (8, 9). Similarly, catabolite activation protein-dependent promoters in E. coli also lack a good −35 sequence (4, 24). By analogy to these systems, one of the two GC-rich inverted repeats found in the φ11 P₁ promoter region could be the binding site for the regulatory protein, Rin. The binding would facilitate RNA polymerase binding at the promoter, thereby activating int expression. However, further experiments are required to localize the Rin binding site in order to confirm this speculation.

The hypothesis that P₁ is the target site for Rin activation was suggested by S1 mapping of the transcription start site of the int gene in the presence or absence of the rin genes. The fact that removal of the P₁ promoter region from plasmid pYL272 resulted in elimination of Xyl/E activity further confirmed this conclusion. These results also exclude the possibility that Rin acted as an antiterminator to terminate transcription from a distal promoter upstream of the int gene.

Our results showed that two genes, rinA and rinB, are required for int gene activation (Fig. 5). Such a requirement is reminiscent of the λ system, in which both CII and CIII are required for the λ int expression initiated from P₁. In λ, specific binding of λ CII near P₁ activates int expression, while CIII stabilizes CII, possibly by protecting it from proteinase degradation (6, 28). Because φ11 RinA is basic and RinB is acidic, RinA may be the DNA-binding protein, equivalent to λ CII, which directly activates int expression. The RinB protein may play a secondary role in activation such as providing stability to RinA, i.e., a role equivalent to that of λ CIII. However, we were unable to find a helix-turn-helix DNA-binding motif, similar to that in the λ CII protein, in RinA.

In contrast to the φ11 system, the activation of L54a int gene transcription required the presence of only the rinA gene. This finding is surprising because the two int genes have almost identical sequences at the potential regulatory sites. The molecular mechanism of the different requirements for activation of the int genes between the two phages is not clear. However, by examining the nucleotide sequences of the L54a and φ11 int genes, we observed that the L54a int gene possessed two translational start sites separated by 142 nucleotides, whereas the φ11 int gene did not (31). We have also shown that the L54a Int protein translated from the second start site is functional. It is possible that the presence of the 142-nucleotide sequence somehow eliminates the need for the RinB function. Further studies are required to clarify this point. The finding that only the rinA gene is required for the activation of L54a int also lends...
support to the notion that RinA may be the major regulatory protein, directly activating int gene expression in φ11.

It is worth noting that the cloned DNA fragment does not contain a promoter for the rin genes. Therefore, a promoter from the plasmid vector must be utilized to transcribe the rin genes, which, in turn, activate int expression. Examination of the recombinant plasmids, such as pYL270 (Fig. 5), suggests that the promoter for transcribing the rin genes is likely to be P1 of the constructed Pint-xylE fusion. If this is the case, how then can P1 be utilized to transcribe the downstream rin genes when its transcription requires transcription of the rin genes? The answer lies in the finding that there is a low level of constitutive expression from P1 which is too low to be detected by the XylE assay. This conclusion is based on the following two observations. (i) In our previous cloning experiments, the int gene was able to mediate recombination when the gene was placed at various sites in different vectors [30, 31]. (ii) The RNA used in the S1 mapping experiment shown in Fig. 2A was obtained from strain RN4220 carrying pCL2112, which contained attP-int-xis but did not contain the rin genes. Thus, in the recombinant plasmids, because of the basal-level transcription from P1, the Rin proteins were initially produced at a low level from P1. The low levels of Rin proteins then positively regulate the transcription from P1 to produce more Rin proteins. Such autogenous activation would continue until an equilibrium state is reached.

Since the activation that we have demonstrated so far was tested only in cis, we also attempted to show that the regulator proteins can also trans activate the P1 promoter. However, all attempts to express the rin genes from several S. aureus promoters failed, possibly because overexpression of rin genes is harmful to S. aureus (as well as to E. coli [32]). Experiments are under way to construct a suitable expression vector for this purpose.

Because expression of the φ11 int gene is regulated by rin at a site within the 110-bp sequence between the translation start sites of the int and xis genes, it is of interest to determine whether the rin genes also regulate xis expression. To this end, we also constructed a Pxis-xylE fusion plasmid (pYL307) similar to that of pYL306 (Fig. 1). Strain RN4220 containing this fusion plasmid exhibited XylE activity at about 6 to 7 mU, indicating that the xis gene is constitutively expressed. When the rin genes were cloned into the fusion plasmid, no change in XylE activity was observed in either orientation, suggesting that the rin genes do not regulate xis expression (results not shown). However, it is possible that Rin is not made in this construct. Since the DNA fragment containing the rin genes in pYL307 is constructed exactly the same as in pCL306 and since XylE activity was observed in pCL307, it is reasonable to assume that in one orientation, the rin genes were expressed from the promoter of the xis-xylE fusion.

The map locations of the rin genes and their target gene, int, on the φ11 map are summarized in Fig. 8. The two loci are separated by about 15 kb, which is about one-third the size of the entire φ11 genome. This region was previously shown to contain the early genes (14, 19). The results presented in this report, therefore, are consistent with these earlier findings.

In conclusion, we have demonstrated that the expression of φ11 int was activated by the rin genes. The activation is most likely at the transcriptional level. The rin locus consists of two genes, rinA and rinB, which are transcribed in the same orientation. One of the genes, rinA, is able to activate the L54α int gene.

FIG. 8. Locations of rinA, rinB, and xis-int-attP on the φ11 map. The EcoRI restriction map of φ11 is shown. Each EcoRI fragment was designated on the basis of restriction size (18, 19). The approximate regions of early, head, and tail genes are shown. Broad arrows represent orientation of transcription. +, activation.

ACKNOWLEDGMENTS

This work was supported in part by research grant DMB-8614082 from the National Science Foundation and by grant BRSG S07 RR05373 awarded by the Biomedical Research Support Grant Program, Division of Research Resources, National Institutes of Health.

REFERENCES

1102 YE AND LEE