Complete Sequence and Enhancer Function of the Homologous DNA Regions of *Autographa californica* Nuclear Polyhedrosis Virus

LINDA A. GUARINO,* MARK A. GONZALEZ, AND MAX D. SUMMERS

Department of Entomology, Texas A&M University, and Texas Agricultural Experiment Station,
College Station, Texas 77843-2475

Received 24 March 1986/Accepted 11 June 1986

The nucleotide sequence of the five regions of homologous DNA in the genome of *Autographa californica* nuclear polyhedrosis virus DNA was determined. The homology of repeated sequences within a region was 65 to 87%, and the consensus sequences for each region were 88% homologous to each other. Sequences proximal to the *EcoRI* sites were most conserved, while the distal sequences were least conserved. The *EcoRI* sites formed the core of a 28-base-pair imperfect inverted repeat. All homologous regions functioned as enhancers in a transient expression assay. A single *EcoRI* minifragment located between *EcoRI*-Q and -L enhanced the expression of 39CAT as efficiently as the regions containing numerous *EcoRI* repeats did.

The genome of the baculovirus *Autographa californica* nuclear polyhedrosis virus (AcNPV) is a circular double-stranded molecule of approximately 128 kilobase pairs. Although the genome consists primarily of unique sequences, five regions which share homologous sequences are interspersed along the length of the genome (6). Because the homologous regions (*hrs*) are rich in *EcoRI* sites, an *EcoRI* restriction digest of AcNPV DNA contains extramolar amounts of small fragments which range in size between 72 and 215 base pairs (bp). A similar pattern of interspersed homologous DNA is also found in another baculovirus, *Choristoneura fumiferana* nuclear polyhedrosis virus (21), suggesting that the *hrs* play an important role in the life cycle of baculoviruses.

The synthesis of viral proteins in AcNPV-infected cells occurs in successive stages, indicating temporal control of viral gene expression (5, 15, 20, 26, 37). At least some of this temporal control is exerted at the level of transcription, as evidenced by the presence of different RNA transcripts at different times of infection (1, 9, 30, 34). Overlapping transcripts have been mapped in several regions of the AcNPV genome (10, 23, 28). Some of these transcripts share a common 5' end, while others are 3' coterminal. It has been hypothesized that this type of transcriptional unit may regulate the expression of baculovirus genes, as the initiation of transcription at upstream promoters may repress initiation of transcription downstream (10). Activation of some of the delayed-early class of promoters may be due to an immediate-early gene which *trans* activates the expression of the 39K gene (17). Activation of the late class of transcripts may be linked to a virus-induced RNA polymerase (12).

Recently, we demonstrated that plasmids containing *hr* DNA stimulate the expression of the chloramphenicol acetyltransferase (CAT) gene under the control of the AcNPV 39K promoter (18). Although this *trans* effect was originally observed as part of our search for *trans*-active factors, we hypothesized that the stimulation of CAT activity was the result of *in vivo* recombination and the resultant *cis* activation. *hr5* was shown to exhibit all the characteristics of an enhancer, including orientation independence and the abilities to function at a distance from the promoter, to activate

heterologous promoters, and to increase RNA polymerase density on the linked gene.

Here we report the nucleotide sequence of the remaining homologous regions and present evidence that they function as enhancers.

MATERIALS AND METHODS

DNA sequencing. The sequencing strategy for the five regions is indicated in Fig. 1. The AcNPV restriction fragments *Pst*I-D (*hr1*), *Hind*III-L (*hr2*), *Hind*III-B (*hr3*), *Bgl*II-G (*hr4*), or *Hind*III-Q (*hr5*) were partially or completely digested with *Eco*RI and ligated with *Eco*RI-digested M13mp9 (27). Single-stranded recombinant phage was purified and sequenced by the dideoxy chain termination procedure (29). The restriction map data generated by the initial sequencing were used to design additional cloning and sequencing to determine the order of the *Eco*RI minifragments. For *hr2*, *Bal* 31 deletions of *Hind*III-L were constructed according to standard procedures (25) and sequenced to determine the order of the minifragments. The sequences were compiled and analyzed by the programs of Devereaux et al. (7).

Analysis of enhancer function. The conditions for cell culture, transfection, and CAT assays have been described previously (14, 17). To determine whether the *hrs* function as enhancers, we repaired the following fragments with the Klenow fragment of DNA polymerase I according to standard procedures (25) and cloned in both orientations into the *Hind*III (Klenow-repaired) site in the multiple cloning site upstream of 39CAT: for *hr1*, the *Cla*I fragment (Fig. 1); for *hr2*, a 1.5-kilobase-pair *Hind*III-*Sal*I fragment; for *hr3*, the *Mlu*I-*Ssp*I fragment (Fig. 1); and for *hr5*, the *Mlu*I fragment (Fig. 1). *hr4left* was cloned (in one orientation with an *Nsi*I-*Xba*I fragment) into the *Pst*I and *Hind*III sites of 39CAT. The resulting plasmid contains *hr4* in an orientation opposite of that in the standard AcNPV genetic map (33).

RESULTS AND DISCUSSION

Nucleotide sequence of AcNPV homologous DNA. The sequencing strategy for the five homologous regions is shown in Fig. 1, and the nucleotide sequences are presented

* Corresponding author.

FIG. 1. Arrangement of homologous sequences on the AcNPV genome (A) and the sequencing strategy of the *hrs* (B). In panel A, the location of the *hrs* in the *EcoRI* restriction map is indicated by heavy vertical lines 1 to 5. The indicated clones which overlap the *hrs* were subcloned for sequencing. Some relevant restriction sites are indicated: C, *Clal*; H, *Hinfl*; Ss, *SspI*; N, *Nsil*; X, *XbaI*; Sa, *Sall*. By convention, the genetic map of AcNPV is presented with the left end of *EcoRI*-I at position 0 (33). In panel B, the indicated restriction fragments were sequenced. In *hr2* Bal 31 deletions were constructed and sequenced. The number of bases between the *EcoRI* sites (E) is indicated.

in Fig. 2. The sequence of *hr5* has been reported elsewhere (18). The sizes of the *EcoRI* minifragments in *hr1*, *hr2*, *hr4*, and *hr5* were previously determined by Cochran and Faulkner (6) by polyacrylamide gel electrophoresis. Our sequence data indicated that their analysis was essentially correct for *hr2* and *hr5*. *hr1* contains four *EcoRI* minifragments, one each of 158 and 89 bp and two of 90 bp. This is one more minifragment than previously reported (6). *hr4* was reported to contain four *EcoRI* fragments between *EcoRI*-Q and -L (6). DNA sequencing showed that only one fragment of 120 bp was located in that site; this is referred to here as

hr4left. The other three fragments (215, 90, and 79 bp) are located between *EcoRI*-L and -E. This region is referred to here as *hr4right*. The size of the *EcoRI* minifragments in *hr3* was not determined by Cochran and Faulkner. However, on the basis of hybridization data, *hr3* was expected to be one of the smaller regions. Sequencing data indicated that *hr3* is the second largest in overall size, with seven *EcoRI* fragments: one each of 148, 111, and 88 bp and four of 72 bp. The poor hybridization of *hr3* in the previous study may be attributable to the fact that *hr3* contains four 72-bp minifragments which are smaller than the probes used for hybridization (6).

FIG. 2. Nucleotide sequence of the *hrs*. The sequences are presented so that they read from left to right on the genetic map (genome orientation).

Analysis of the *hr* nucleotide sequence. The homology of the repeated sequences in *EcoRI* minifragments is demonstrated in Fig. 3. In this analysis, gaps were inserted for optimal alignment of homologous sequences. Much of the sequence heterogeneity within each region can be attributed to the

additional bases in the longer minifragments. The sequences surrounding the *EcoRI* sites were the most conserved, and the sequences distal to the *EcoRI* sites were the least conserved.

Homologous sequences extending to the left and right of

FIG. 3. Alignment of homologous sequences. The nucleotide sequences of the five *hrs* are presented with gaps (.) inserted for optimal alignment of homologous sequences. Some bases from the longer fragments are not shown; the number of bases is indicated in parentheses. *Eco*RI restriction sites are underlined. The numbers on the left refer to the base numbers in Fig. 2. con, Consensus sequence for each region; —, where the consensus can be more than 1 base.

the first and last *Eco*RI sites were also included. The number of homologous bases beyond the *Eco*RI sites was variable in both directions, and there was apparently no sequence common to all 5' or 3' ends of the *hrs*.

The percent homology between the consensus sequences for each region and the individual *Eco*RI minifragments was 80% for *hr1*, *hr2*, and *hr3*; 65% for *hr4right*; and 87% for *hr5* in Table 1. By this calculation, *hr4right* showed the least

homology within a single region. This was a reflection of the large difference in size between the 215-bp minifragment and the two smaller fragments.

The homology between the consensus sequences of each region is shown in Fig. 4. The homology of the consensus sequences was 88% when two regions (*hr1* and *hr3*) were compared in the genome antisense orientation. The homology between the consensus sequences was only 75% (not

FIG. 4. Alignment of consensus sequences for each *hr*. In some cases the dashes have been changed to the base which best fits the overall consensus. The sequences for *hr1* and *hr3* are presented in the opposite (—) orientation as in the standard AcNPV genetic map; the other *hrs* are shown in the standard (+) orientation. ., Gap in the known sequence.

- DNA sequences interspersed at five regions in the baculovirus AcMNPV genome. *J. Virol.* **45**:961-970.
7. Devereaux, J., P. Haerberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**:387-396.
 8. de Villiers, J., and W. Schaffner. 1981. A small segment of polyoma virus DNA enhances the expression of a cloned B-globin gene over a distance of 1400 base pairs. *Nucleic Acids Res.* **9**:6251-6264.
 9. Esche, H., H. Lübbert, B. Siegmann, and W. Doerfler. 1982. The translation map of the *Autographa californica* nuclear polyhedrosis virus AcNPV genome. *EMBO J.* **1**:1629-1633.
 10. Friesen, P. D., and L. K. Miller. 1985. Temporal regulation of baculovirus RNA: overlapping early and late transcripts. *J. Virol.* **54**:392-400.
 11. Fromm, M., and P. Berg. 1982. Deletion mapping of DNA required for SV40 early region promoter function in vivo. *J. Mol. Appl. Genet.* **1**:457-481.
 12. Fuchs, L. Y., M. S. Woods, and R. F. Weaver. 1983. Viral transcription during *Autographa californica* nuclear polyhedrosis virus infection: a novel RNA polymerase induced in infected *Spodoptera frugiperda* cells. *J. Virol.* **48**:641-646.
 13. Gluzman, Y., and T. Shenk (ed.). 1983. Enhancers and eukaryotic gene expression. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 14. Gordon, J. D., and E. B. Carstens. 1984. Phenotypic characterization and physical mapping of a temperature-sensitive mutant of *Autographa californica* nuclear polyhedrosis virus defective in DNA synthesis. *Virology* **138**:69-81.
 15. Gorman, C. M., L. F. Moffat, and B. H. Howard. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* **2**:1044-1051.
 16. Gruss, P., R. Dhar, and G. Khoury. 1981. Simian virus 40 repeated sequences as an element of the early promoter. *Proc. Natl. Acad. Sci. USA* **78**:943-947.
 17. Guarino, L. A., and M. D. Summers. 1986. Functional mapping of a *trans*-activating gene required for expression of a baculovirus delayed-early gene. *J. Virol.* **57**:563-571.
 18. Guarino, L. A., and M. D. Summers. 1986. Interspersed homologous DNA of *Autographa californica* nuclear polyhedrosis virus enhances delayed-early gene expression. *J. Virol.* **60**:215-223.
 19. Hearing, P., and T. Shenk. 1983. The adenovirus type 5 E1A transcriptional control region contains a duplicated enhancer element. *Cell* **33**:695-703.
 20. Kelly, D. C., and T. Lescott. 1981. Baculovirus replication: protein synthesis in *Spodoptera frugiperda* cells infected with *Trichoplusia ni* nuclear polyhedrosis virus. *Microbiologica* **4**:35-47.
 21. Kuzio, J., and P. Faulkner. 1984. Regions of repeated DNA in the genome of *Choristoneura fumiferana* nuclear polyhedrosis virus. *Virology* **139**:185-188.
 22. Lang, J. C., D. A. Spandidos, and N. M. Wilkie. 1984. Transcriptional regulation of a herpes simplex virus immediate early gene is mediated through an enhancer-type sequence. *EMBO J.* **3**:389-395.
 23. Lübbert, H., and W. Doerfler. 1984. Transcription of overlapping sets of RNAs from the genome of *Autographa californica* nuclear polyhedrosis virus: a novel method for mapping RNAs. *J. Virol.* **52**:255-265.
 24. Lusky, M., L. Berg, H. Weiher, and M. Botchan. 1983. Bovine papilloma virus contains an activator of gene expression at the distal end of the early transcription unit. *Mol. Cell. Biol.* **3**:1108-1122.
 25. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 26. Maruniak, J. E., and M. D. Summers. 1981. *Autographa californica* nuclear polyhedrosis virus phosphoproteins and synthesis of intracellular proteins after virus infection. *Virology* **109**:25-34.
 27. Messing, J. 1983. New m13 vectors for cloning. *Methods Enzymol.* **101**:20-78.
 28. Rankin, C. B., B. F. Ladin, and R. F. Weaver. 1986. Physical mapping of temporally regulated, overlapping transcripts in the region of the 10K protein gene in *Autographa californica* nuclear polyhedrosis virus. *J. Virol.* **57**:18-27.
 29. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463-5468.
 30. Smith, G. E., J. M. Vlak, and M. D. Summers. 1982. In vitro translation of *Autographa californica* nuclear polyhedrosis virus early and late mRNAs. *J. Virol.* **44**:199-208.
 31. Tyndall, C., G. La Mantia, C. M. Thacker, J. Favoloro, and R. Kamen. 1981. A region of the polyoma virus genome between the replication origin and late protein coding sequences is required in *cis* for both early gene expression and viral DNA replication. *Nucleic Acids Res.* **9**:6231-6250.
 32. Vlak, J. M., and K. G. Odink. 1979. Characterization of *Autographa californica* nuclear polyhedrosis virus deoxyribonucleic acid. *J. Gen. Virol.* **44**:333-347.
 33. Vlak, J. M., and G. E. Smith. 1982. Orientation of the genome of *Autographa californica* nuclear polyhedrosis virus: a proposal. *J. Virol.* **41**:1118-1121.
 34. Vlak, J. M., G. E. Smith, and M. D. Summers. 1981. Hybridization selection and in vitro translation of *Autographa californica* nuclear polyhedrosis virus mRNA. *J. Virol.* **40**:762-771.
 35. Weeks, D. L., and N. C. Jones. 1983. E1A control of gene expression is mediated by sequences 5' to the transcriptional starts of the early viral genes. *Mol. Cell. Biol.* **3**:1222-1234.
 36. Weiher, H., M. König, and P. Gruss. 1983. Multiple point mutations affecting the simian virus 40 enhancer. *Science* **219**:626-631.
 37. Wood, H. A. 1980. *Autographa californica* nuclear polyhedrosis virus-induced proteins in tissue culture. *Virology* **102**:21-27.