Binding of the herpes simplex virus major regulatory protein to viral DNA

(Fab antibody fragments/methylation interference/gel retardation)

NANCY MICHAEL AND BERNARD ROIZMAN

The Marjorie B. Kovler Viral Oncology Laboratories, The University of Chicago, Chicago IL 60637

Contributed by Bernard Roizman, October 3, 1989

ABSTRACT Infected-cell protein 4 (ICP4), the major regulatory protein specified by herpes simplex virus 1 in infected cells, binds to homologs of the sequence ATCGTCnnnnYC-GRC (A sites, where n is any nucleotide, Y is a pyrimidine, and R is a purine) and to unrelated sequences for which no consensus sequence has been derived (B sites). We have examined the binding of ICP4 to each of two A and two B binding sites by using Fab fragments of a monoclonal antibody that is reactive with an epitope located at the N terminus of ICP4 and that decreases the mobility of ICP4–DNA complexes in nondenaturing gels. The results indicate that each type of site binds two monomers of ICP4. Methylation-interference studies on the type B sites mapped the guanines whose methylation interfered with the binding of ICP4. The methylation-interference pattern obtained with one of the B sites was similar to that obtained on an A site but differed from that of the other B site. The ability of ICP4 to bind to DNA fragments containing the binding site appears to be dependent on length and on the proximity of the binding site to the fragment end. Short DNA fragments did not form stable complexes with ICP4 even though they contained all of the purines whose methylation interfered with the binding of the regulatory protein.

Infected-cell protein 4 (ICP4) is the major regulatory protein of herpes simplex virus 1 (HSV-1); it regulates HSV-1 genes both positively and negatively (for review, see ref. 1). Unambiguous demonstration that ICP4 binds directly to viral DNA sequences has come from two sources. Michael *et al.* (2) demonstrated that ICP4 contained in nuclear extracts of infected cells that had been electrophoretically separated in denaturing gels, electrically transferred to nitrocellulose, and renatured specifically bound HSV-1 DNA fragments containing ICP4 binding sites. Kattar-Cooley and Wilcox (3) demonstrated that preparations of ICP4 lacking impurities, as determined by Coomassie blue-stained gels containing electrophoretically separated proteins, also bound a DNA fragment with a known binding site. However, three puzzles remain to be solved.

The first puzzle concerns the sequence recognized by ICP4. Faber and Wilcox (4) compared the sequences of a HSV-1 DNA fragment with the sequences of two fragments from pBR322 that bound ICP4 and derived the consensus ATCGTCnnnnYCGRG (where n is any nucleotide, Y is a pyrimidine, and R is a purine) for the binding site. One of the sites mapped by Kristie and Roizman (5) conformed with the derived consensus whereas the second did not. Although several laboratories reported that the nonconsensus sites either were degenerate binding sites or did not exist (6–8), there is little doubt that binding sites which do not conform to the consensus derived by Faber and Wilcox (4) do exist and are present in both 5' untranscribed and transcribed noncod-

ing domains of many genes (2, 9, 10). Both consensus and nonconsensus sites bind to ICP4 that had been electrophoretically separated in denaturing gels, electrically transferred to nitrocellulose, and renatured (2). Inasmuch as there is no basis for differentiation of consensus and nonconsensus sites functionally (refs. 8 and 10–12; M. G. Romanelli, P. Mavromara-Nazos, and B.R., unpublished results), it is convenient to refer to these sites as A and B, respectively. Does ICP4 bind to two or more distinct sequences or does the protein recognize a common motif that is not presently perceived?

The second puzzle concerns the molecular nature of the interaction of ICP4 with DNA. ICP4 has been reported to exist predominantly as a homodimer (13) but also possibly as a tetramer (3) and, therefore, the question arises whether ICP4 binds to DNA as a dimer. The question is relevant since neither the consensus sequence (representative of A sites) nor the B sites exhibit credible palindromes or repeats that would be recognized by two identical molecules in a similar manner.

Lastly, in the course of analyses of HSV-1 DNA fragments for binding sites, Michael *et al.* (2) noted that several fragments formed more than one ICP4-specific band in nondenaturing gels. The question is whether these complexes represent a higher order of binding of ICP4 to DNA or a single ICP4–DNA complex to which other proteins are bound as well.

To resolve at least some of these issues, we used methylation interference to map the guanines likely to make contact with ICP4 in DNA probes and we took advantage of the technique reported by Kristie and Roizman (9) in which antibody to the DNA-binding protein is used to retard the migration of the complex in nondenaturing gels. In this paper we report that two monomers of ICP4 bind to one binding site, either type A or type B, and, therefore, A and B sites do not differ with respect to the number of ICP4 molecules bound. Our results also suggest a similarity in the methylation-interference patterns obtained on one A site and one B site.

MATERIALS AND METHODS

Cells, Virus, and Protein Extracts. HeLa cells, grown to confluence in 150-cm² flasks, were mock-infected or infected with 5 plaque-forming units of HSV-1 strain F [HSV-1 (F)] per cell and maintained for 18 hr. Nuclear extracts (0.42 M) were prepared as described (5) and frozen without dialysis at –80°C.

DNA–Protein Binding Assays. The DNA probes and their location in the viral genome are shown in Fig. 1. DNA–protein binding assays were done as follows: 0.5 ng of probe DNA, 0.5 μg of the 0.42 M nuclear extract, 0.5 μg of poly(dI·dC), or 0.1 μg of poly(dI·dC) plus 0.4 μg of poly(dA·dT)·poly(dA·dT) were incubated in 12 μl of buffer A [20

FIG. 1. Sequence arrangements and location of the probe DNAs in the genome of HSV-1. (A) Location of the *Bam*HI fragments S and N and the relative location of the $\alpha 0$, $\alpha 4-1$ and $\alpha 4-2$ probes in HSV-1 DNA. The boxes represent the repeats ab and $a'b'$, flanking the unique sequences of the long component, and $c'a'$ and ac , flanking the unique sequences of the short component of HSV-1 DNA. The arrows represent the location of the start site and the direction of transcription of the $\alpha 0$ and $\alpha 4$ genes. (B) Schematic diagram showing the nucleotide sequence, fragment length, and composition of the DNA probes used for the studies of the B type ICP4 binding sites in the $\alpha 4-2$ site. The nucleotide sequence of the HSV-1 59-bp fragment containing the nucleotides -194 to -135 from the 5' untranscribed domains of the $\alpha 4$ gene is shown in large letters at the top of the diagram; the small letters show the adjacent nucleotide sequences of the cloning vector present in some of the probes. Lines A-H indicate the sequences present in the various probes used in these studies. The thick and thin lines represent HSV-1 and vector sequences, respectively, present in the probe DNAs. The numbers are the nucleotide numbers at the ends of the HSV-1 sequences. The DNA probes were as follows: $\alpha 0$, nucleotides -110 to +72 of the $\alpha 0$ gene cloned as pRB3563 (9); $\alpha 4-2$, nucleotides -47 to +33 of the $\alpha 4$ gene cloned as pRB2609 (2); A, nucleotides -194 to -135 of the $\alpha 4$ gene cloned as pRB3085 (5); B, nucleotides -194 to -174 and -155 to -135 of the $\alpha 4$ gene cloned as pRB3188 (5); C, nucleotides -181 to -135 of the $\alpha 4$ gene cloned as pRB3191 (5); D, nucleotides -157 to -135 obtained by cleavage of pRB3085 with *Eco*RI, *Hind*III, and *Sac* II; E, nucleotides -194 to -158 obtained by cleavage of pRB3058 with *Eco*RI, *Hind*III, and *Sac* II; F, nucleotides -194 to -147 obtained by cleavage of pRB3058 with *Eco*RI, *Hind*III, and *Apa* I; G, nucleotides -196 to -172 of the $\alpha 4$ gene (each strand of this fragment was made in an Applied Biosystems DNA synthesizer; the two strands were annealed and the double-stranded fragment was purified by gel electrophoresis); H, nucleotides -146 to -135 obtained by cleavage of pRB3085 with *Eco*RI, *Hind*III, and *Apa* I.

mM Tris-HCl, pH 7.6/50 mM KCl/1 mM EDTA/0.05% Nonidet P-40/5% (vol/vol) glycerol/5 mM 2-mercaptoethanol/bovine serum albumin (200 μ g/ml)] for 20 min at room temperature. The reaction mixtures were electrophoretically separated in 4% polyacrylamide gels as described (ref. 5 and references therein).

Fab Preparation and Antibody Titration Assay. Monoclonal antibody H943, obtained from Lenore Pereira (University of California, San Francisco), was shown to react with the epitope formed by the amino acids 120-135 of the HSV-1 (14) and to retard the migration of ICP4-DNA complexes in nondenaturing gels (5). Monovalent Fab fragments were prepared from H943 IgG derived from ascites fluid by cleavage of the IgG molecule with papain (ImmunoPure Fab preparation kit 44885; Pierce) used as specified by the manufacturer. Dilutions of Fab fragments, IgG molecules, or of total ascites fluid protein were made in buffer A immediately prior to use. The diluted antibody (2 μ l) was added to a DNA-protein mixture described above and allowed to react for an additional 20 min at room temperature before electrophoresis.

RESULTS

A Sites in the $\alpha 0$ and $\alpha 4$ Genes Each Bind Two Molecules of ICP4. The objective of these experiments was to determine the number of molecules of ICP4 that bind to two A sites on HSV-1 DNA (Fig. 1). The experimental design was based on the observation that the binding of monoclonal anti-ICP4 antibody to the ICP4-DNA complex retards the migration of the complex in nondenaturing gels (5, 9). In the presence of

an excess of ICP4-DNA complexes relative to antibody molecules, at least a fraction of the ICP4-DNA complexes should bind only one molecule of IgG. As the relative amount of antibody molecules increases, more ICP4-DNA complexes would approach saturation binding by the antibody. Under these conditions we would expect that the electrophoretic mobility of the complexes in the nondenaturing gels would be inversely related to the number of antibody molecules bound by the complexes. To avoid complications resulting from two ICP4-DNA complexes being bound by one molecule of IgG, all of the experiments were done with univalent Fab fragments derived by papain digestion of monoclonal antibody H943 specific for an epitope located between amino acids 120 and 135 in ICP4 (14).

Fig. 2 shows the results of antibody titration assays on the two A binding sites, the $\alpha 0$ site mapped by Kristie and Roizman (5) in the 5' untranscribed domain of the $\alpha 0$ gene and the $\alpha 4-2$ site mapped by Muller (6) and by Faber and Wilcox (7), which extends across the transcription initiation site of the $\alpha 4$ gene. In the absence of the antibody, the ICP4-DNA complex formed only one band. At low concentration of Fab fragments, the complex formed two bands, one comigrating with the original band and one migrating slightly slower. As the concentration of the antibody was increased, the original band disappeared and a third band migrating more slowly became dominant. At highest concentrations of antibody, only the slowest migrating ICP4-DNA-Fab band was detected. This band migrated more rapidly than the complex consisting of ICP4-DNA-IgG, as would be expected since the mass of a single IgG molecule is greater than that of two Fab molecules. Because the ICP4-DNA complex formed two

FIG. 2. Fab titration assay on type A binding sites. Autoradiographic image of a 4% polyacrylamide gel containing electrophoretically separated ICP4–DNA–Fab complexes. The sources of the probe DNAs are shown in Fig. 1A. DNA fragments were 5'-end-labeled with [γ - 32 P]ATP by T4 polynucleotide kinase to an average activity of 50,000 cpm/ng of fragment DNA. Competitor DNA (Pharmacia) was prepared according to manufacturer's recommendations as a stock at 1.0 μ g/ μ l. The reaction mixture contained of 0.5 μ g of poly(dI-dC) and no (0) antibody, IgG, or 2 μ l of a 1:2 dilution series of the Fab monoclonal antibody fragments; the calculated amount of undiluted Fab fragments added is indicated above each lane (in μ l). N, the numbers 0, 1, or 2 at the right indicate DNA–ICP4 complexes with 0, 1, or 2 Fab antibody fragments bound, respectively.

distinct bands whose migration and abundance were determined by the concentration of Fab antibody in the reaction mixture, we conclude that two monomers of ICP4 are contained in each ICP4–DNA complex.

Analyses by Methylation Interference of the Binding of ICP4 to the B Site in the $\alpha 4$ Gene. The B binding site chosen for these studies was the sequence located by Kristie and Roizman (5, 9) in the 5' untranscribed domain of the $\alpha 4$ gene and shown to bind ICP4. This 59-base-pair (bp) sequence was chosen for three reasons. (i) It lacked a homolog of the consensus sequence. (ii) Deletion analyses done in the course of their studies (5) failed to pinpoint the exact binding site and led to the suspicion that there may be more than one. (iii) The complexes formed by ICP4 and DNA fragments containing the 59-bp HSV-1 sequence consistently formed two or more bands in nondenaturing gels.

To analyze the complex, it was necessary to map the binding site(s) more precisely. The results of the methylation-interference studies are shown in Fig. 3. The significant feature of the results is that there appeared to be two sites in which methylation of guanines interfered with the binding of ICP4 to the HSV-1 sequence. The two sites are the more cap-site distal (5' coding) GGGcgcGccG and the cap-site proximal (5' coding) GGcgcGcG sequences. (The guanines whose methylation interfered with the binding of ICP4 are shown as underlined capital letters.)

Similar studies were done on the ICP4 binding site mapped by Kristie and Roizman (5) in the 5' untranscribed domain of

FIG. 3. Methylation-interference footprint of ICP4 on probe A. Autoradiographic image of the products recovered after piperidine cleavage of partially methylated DNA electrophoretically separated on an 8% polyacrylamide/urea sequencing gel. Probe DNA, end-labeled at the 5' terminus of the coding or noncoding strand, was methylated as described (5). The modified probe was reacted in a standard DNA binding assay. The ICP4 bound and free DNA, resolved on a 4% polyacrylamide gel, were recovered from a gel slice by electroelution and cleaved with piperidine, and the cleavage products were electrophoretically separated on an 8% polyacrylamide/urea sequencing gel as described. (15). Lanes: G or G+A, Maxam and Gilbert cleavage reactions of the partially methylated probe A DNA; FREE, cleavage pattern of protein free DNA; BOUND, cleavage pattern of the DNA found in the ICP4–DNA complex. The sequence of the relevant portions of the probe DNA is indicated. Guanines under represented in the cleavage pattern are indicated by a star.

the $\alpha 0$ gene. In this instance, the guanines whose methylation interfered with the binding of ICP4 were contained in the

sequence atcGtcactGccG (data not shown). In this instance the guanines making contact with the DNA were within a homolog of the consensus sequence.

Two Molecules of ICP4 Bind to Each of the Two B Sites in the $\alpha 4$ Gene. The derivation and domains of the series of DNA fragments from the 59-bp HSV-1 (F) sequence, nucleotides –194 to –135 from the 5' untranscribed domain of the $\alpha 4$ gene, are shown in Fig. 1B. The results of analyses of the binding of ICP4 to these fragments are shown in Fig. 4 and may be summarized as follows.

FIG. 4. Fab titration assay on type B binding sites. Autoradiographic image of a 4% polyacrylamide gel containing electrophoretically separated ICP4–DNA–Fab complexes. The assays were done as described in Fig. 2 except that the competitor DNA was poly(dAdT)·poly(dAdT) (0.4 μ g) and poly(dIdC) (0.1 μ g). The calculated amount of undiluted Fab fragment added is indicated above each lane (in μ l). (A–F) DNA fragments A–F, respectively, were used and are shown in Fig. 1B.

(i) The ICP4–DNA fragment A complex formed several bands (Fig. 4A) in nondenaturing gels. Two slowly migrating bands were formed in the presence of the highest concentration of Fab fragments tested. At low Fab concentrations, there appeared to be two sets of bands. The fastest migrating set was similar to those obtained with the A sites in the $\alpha 0$ and $\alpha 4$ genes shown in Fig. 2. In this set the ICP4–DNA complexes formed three bands; with increasing concentration of Fab molecules, the complexes shifted progressively from the band with the fastest mobility to that of the slowest mobility. Although a similar trend was seen with the slower migrating set of complexes, the number of bands was less readily discernible although clearly there were more than three bands.

(ii) A more detailed picture of the binding of ICP4 to different domains of the 59-bp HSV-1 sequence is apparent in Fig. 4 B–F. Fragment B contains both sets of purines whose methylation interferes with the binding of ICP4, but much of the intervening sequence (nucleotides –173 to –157) has been deleted. This fragment formed one fast migrating set of three strong bands and one set of barely detectable bands with ICP4 and the Fab fragments (Fig. 4B). Fragment C, D, E, or F formed a single ICP4-specific band in the absence of antibody, and two additional more slowly migrating bands at low and high Fab concentrations. Since fragments D and E bound ICP4, it follows that the downstream site does not extend upstream past nucleotide –157, whereas the upstream site does not extend downstream past nucleotide –174. Since ICP4–DNA fragment complexes for each of the sites formed two additional bands in the presence of Fab fragments, each site bound two monomers of ICP4.

(iii) The synthetic oligonucleotide containing the nucleotides –196 to –172 (fragment G) did not form a detectable complex with ICP4 (data not shown). The methylation-interference results suggest that the nucleotides likely to be directly involved in binding ICP4 are located between nucleotides –189 and –179. These sequences are centrally positioned on this oligonucleotide; nevertheless this sequence failed to form a stable complex with ICP4.

(iv) The sequences required for stable binding of ICP4 to the binding site proximal to the transcription initiation site may extend upstream from the nucleotide –146. This conclusion is based on the observation that in contrast to

fragment D (Fig. 4D), fragment H containing the HSV-1 nucleotides –135 to –146 did not bind ICP4 (data not shown).

DISCUSSION

Number of Molecules of ICP4 Binding to the A and B Binding Sites. The experimental design of the studies reported in this paper was based on the observation that monoclonal antibodies to DNA binding proteins can decrease the electrophoretic mobility of DNA–protein complexes in nondenaturing gels (5, 9, 16). It could be expected, therefore, that in the presence of Fab fragments of a specific monoclonal antibody, if more than one molecule of ICP4 were bound at a site, the ICP4–DNA complex would form a ladder of bands decreasing in electrophoretic mobility as a function of saturation of the epitopes recognized by the antibody in ICP4. The number of rungs in the ladder, less the one formed in the absence of antibody, would correspond to the number of ICP4 monomers bound to each site. We report studies on two A sites and two B sites present in close proximity in one DNA fragment. These studies indicated that each A and B ICP4 binding site bound two monomers of ICP4, as determined from the number of ICP4–DNA bands formed in nondenaturing gels in the absence of antibody and as a function of increasing amounts of antibody. Similar results were obtained with a type B binding site (2) present in the domain of late α (γ) gene (data not shown). We conclude (i) that the differences between the A and B binding sites for ICP4 cannot be attributed to the number of ICP4 molecules bound and (ii) Fab fragments of monoclonal antibodies to DNA binding proteins are a powerful tool for determining the number of identical protein molecules binding to DNA at specific sites.

We should note parenthetically that we cannot exclude the possibility that stable monomeric ICP4–DNA complexes could form. In some studies we have noted weak, more rapidly migrating bands containing ICP4 as determined by their reactivity with monoclonal antibody. We should also note that DiDonato and Muller (17) reported that ICP4– $\alpha 4$ -2 DNA complexes formed multiple bands at high concentrations of ICP4. Although we cannot exclude the possibility that more than two molecules of ICP4 can bind to a single binding site, slow migrating complexes could result from aggregation of ICP4 as a consequence of partial denaturation.

Studies on B Type Binding Sites. The demonstration that the 59-bp $\alpha 4$ -1 site contains two ICP4 binding sites explains earlier difficulties in localizing precisely these sites by deletion analyses (5). Although there is little homology between the nucleotide sequences of the two B sites or between the two B sites and authenticated A sites, similarities do exist between some of the methylation-interference patterns obtained to date. Thus the pattern of guanines whose methylation interferes with the binding of ICP4 to DNA in $\alpha 4$ -1 distal sites GGnnnnnGnnG is similar to the $\alpha 0$ methylation inter-

ference pattern nnnGnnnnnGnnG nnnnnGnGnnGnn. The proximal $\alpha 4$ -1 site nnnnnGnGnnn GGnnnnnnnnGnn is either a degenerate binding site similar to the distal site or an entirely different site. A precise definition of the nature of the differences among the ICP4 binding sites requires mapping of additional binding sites.

We should note that experimental detection of a binding site must take into account the context of the sequence in which the binding sites are imbedded. Thus both proximal and distal $\alpha 4$ -1 binding sites and adjacent nucleotides are very G+C-rich and are likely to be sensitive to competing DNA molecules of similar base composition. The A sites, especially the $\alpha 4$ -2 and the $\alpha 0$ sites, are much less G+C-rich. The difference in the type of DNA competitor used may well explain the failure of other laboratories to detect the $\alpha 4$ -1 sites (6-8).

DNA Sequence Requirements for Binding ICP4. In two instances reported above, we failed to obtain binding of ICP4 to small DNA fragments (fragments G and H; Fig. 1) that contain a binding site as determined by the methylation-interference studies on DNA probe fragment A. We have no data that bears on these observations. Conceivably either the DNA sequence making contact with ICP4 must be contained within the context of a specific nucleotide sequence or the binding site must be at a certain distance from the terminus of the DNA fragment probe. Should the proximity of the binding site to the terminus of the DNA fragment be a

significant factor, the results would suggest that (i) ICP4 alters the conformation of the DNA at the binding site, (ii) in the process the DNA strands in a short fragment become denatured, and (iii) the ICP4-DNA complex becomes unstable.

We thank Lenore Pereira for the gift of the H943 monoclonal antibody. These studies were aided by grants from the National Cancer Institute (CA47451) and the National Institute for Allergy and Infectious Diseases (AI124009 and AI1588-11). N.M. is a predoctoral trainee supported by a U.S. Public Health Service grant from the National Cancer Institute (5 T32 CA09241-10).

1. Roizman, B. & Sears, A. M. (1989) in *Virology*, eds. Fields, B. & Knipe, D. (Raven, New York), 2nd Ed., pp. 1795-1841.
2. Michael, N., Spector, D., Mavromara-Nazos, P., Kristie, T. M. & Roizman, B. (1988) *Science* **239**, 1531-1534.
3. Kattar-Cooley, P. & Wilcox, K. W. (1980) *J. Virol.* **63**, 696-704.
4. Faber, S. W. & Wilcox, K. W. (1986) *Nucleic Acids Res.* **14**, 6067-6083.
5. Kristie, T. M. & Roizman, B. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 4700-4704.
6. Muller, M. T. (1987) *J. Virol.* **61**, 858-865.
7. Faber, S. W. & Wilcox, K. W. (1988) *Nucleic Acids Res.* **16**, 555-570.
8. Roberts, M. S., Boundy, A., O'Hare, P., Pizzorno, M. C., Ciuffo, D. M. & Hayward, G. S. (1988) *J. Virol.* **62**, 4307-4320.
9. Kristie, T. M. & Roizman, B. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 3218-3222.
10. Tedder, D. G., Everett, R. D., Wilcox, K. W., Beard, P. & Pizer, L. I. (1989) *J. Virol.* **63**, 2510-2520.
11. Kristie, T. M. & Roizman, B. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 4065-4069.
12. Resnick, J., Boyd, B. A. & Haffey, M. L. (1989) *J. Virol.* **63**, 2497-2503.
13. Metzler, D. W. & Wilcox, K. W. (1985) *J. Virol.* **55**, 329-337.
14. Hubenthal-Voss, J., Houghten, R. A., Pereira, L. & Roizman, B. (1988) *J. Virol.* **62**, 454-462.
15. Kristie, T. M. & Roizman, B. (1988) *J. Virol.* **62**, 1145-1157.
16. McKnight, J. L. C., Kristie, T. M. & Roizman, B. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 7061-7065.
17. DiDonato, J. A. & Muller, M. T. (1989) *J. Virol.* **63**, 3737-3747.