Nucleic Acids Research, 1994, Vol. 22, No. 16 3331–3338

Minor groove contacts are essential for an interaction of the human cytomegalovirus IE2 protein with its DNA target

Dieter Lang and Thomas Stamminger*
Institut für Klinische und Molekulare Virologie der Universität Erlangen-Nürnberg, Schloßgarten 4, 91054 Erlangen, Germany

Received June 13, 1994; Revised and Accepted July 26, 1994

ABSTRACT

The 86 kDa immediate early-2 protein (IE2, IE86) of human cytomegalovirus (HCMV) is a multifunctional polypeptide that can regulate gene expression both positively and negatively. In particular, it represses its own mRNA synthesis by binding directly to a sequence element, termed *cis* repression signal (CRS), that is located between the TATA box and the transcriptional start site of the major IE enhancer/promoter of HCMV. Here, we provide evidence that IE86, unlike most sequence-specific DNA-binding proteins, interacts primarily within the minor groove of the DNA helix. This was shown by hydroxyl radical and methylation interference assays. In addition, binding studies with inosine-substituted oligonucleotides which have an altered major groove morphology without changing the surface of the minor groove, confirmed the results obtained in interference analyses. This establishes IE86 as a member of a small group of DNA binding proteins that interact with A – T rich sequences within the minor groove and which also includes the TATA-box binding protein TBP. Remarkably, IE86 and TBP are able to bind simultaneously in an immediate vicinity at the major IE enhancer/promoter of HCMV. As minor groove binding proteins are known to bend DNA heavily this could contribute to the observed negative regulation of transcription by IE86.

INTRODUCTION

Human cytomegalovirus (HCMV), one of the seven human herpesviruses, is of considerable clinical importance in newborns and immunosuppressed patients. For instance, it is estimated that approximately 90% of patients with acquired immunodeficiency syndrome (AIDS) will manifest CMV infection, and, of those, 25% will proceed to some form of sight- or life-threatening disease such as retinitis or pneumonia (1).

The double-stranded DNA genome of HCMV comprises 229 345 nucleotides and has the potential to code for more than 200 proteins (2). As shown for all members of the family of herpesviruses investigated so far, lytic cycle gene expression of HCMV occurs in a sequential fashion. After entry into the cell, immediate early (IE) genes are the first to be expressed, followed by the early (E) and late (L) genes (3–5). Immediate-early genes mainly code for regulatory proteins that are involved in the switch from restricted to extensive expression of the genome. Thus, these proteins are proposed to play a key role in determining whether the virus enters lytic cycle gene expression or remains in a latent state within the infected cell. Although regulatory activities have been described for several IE-proteins of HCMV (6–8), the mechanisms used by these polypeptides are largely unknown. The best characterized regulatory protein of HCMV is the 86 kDa IE2 (IE86) polypeptide (9–11). The encoding RNA of 2.2 kb is expressed under control of the major IE enhancer/promoter of HCMV (12,13) and arises via differential splicing (14).

Like for the E1a protein of adenovirus, IE86 of HCMV has been described as a promiscuous transactivator of viral, as well as cellular, gene expression (15–17). For some regulatory sequences like the hsp70 promoter and the HIV LTR it could be demonstrated that a minimal, TATA-box containing promoter is sufficient to mediate a response to IE86 (18). This suggests that the basal transcription apparatus is one target of IE86 effects and correlates with the observation that IE86 is able to interact with the TATA-binding protein TBP (18) and with TFIIB (19). However, IE86 transactivation of some HCMV early genes is stimulated significantly by the presence of regulatory sequences upstream of the TATA box (20–22). In part, this may be mediated by protein/protein contacts as IE86 appears to be able to interact with a large number of as yet unidentified cellular proteins (23).

In addition to its function as transactivator, IE86 has been shown to negatively autoregulate its own expression; it is able to repress transcription from the major IE enhancer/promoter of HCMV (16,24). This negative regulation is dependent on a sequence element, termed the *cis* repression signal (CRS), that is located between the TATA box and the transcriptional start site of the major IE enhancer-promoter (see Figure 1) (25–27). The CRS element which has a partially palindromic structure can confer an IE86 dependent repression to heterologous promoters in an orientation independent but location dependent manner (25–27). While the CRS has been defined by deletion and mutational analysis some years ago, it was only recently possible to demonstrate a direct binding of the IE86 protein to sequences

*To whom correspondence should be addressed

of the CRS element (28–30). In part, this was due to an unusual characteristic of IE86: its interaction with DNA is sensitive to certain non-specific competitor DNAs such as poly(dIdC) (28,30), a property shared with the TATA-binding protein TBP. In addition, the detection of several binding sites for IE86 within an early promoter of HCMV revealed a weak similarity of the respective sequences, suggesting a rather loose sequence requirement for specific recognition of the target DNA (20).

This study was therefore undertaken in order to obtain a detailed picture of the interaction characteristics of IE86 with DNA. We show that IE86 contacts DNA primarily via the minor groove, a property shared with a limited number of other transcription factors such as the TATA binding protein TBP (31–34) or the HMG domain protein LEF1 (35). Remarkably, IE86 and TBP which are both minor groove binding proteins are able to interact simultaneously in an immediate vicinity at the major IE enhancer/promoter of HCMV (36). As minor groove binding proteins are known to bend DNA heavily this could contribute to the observed negative regulation of transcription by IE86.

MATERIALS AND METHODS

Oligonucleotides

Oligonucleotides were obtained from MWG Biotech (Ebersberg, Germany) and Eurogentec (Seraing, Belgium). The following oligonucleotides (5′ to 3′ sequences; double-stranded oligonucleotides are indicated by double shills) were used for gel retardation experiments and methylation interference analysis:

CRS, CGTTTAGTGAACCGTCAGAT(CRS-1)//
CTAGATCTGACGGTTCACTAAACGAGCT (CRS-2);
CRS-MUT (mutated CRS), GCGGCGGTGAACCGTCAGAT//
CTAGATCTGACGGTTCACCGCCGCAGCT;
EAIE2, CTCGAGAAAGTAGCGTTGCGATTTGCAGTCCGCTCAAGCT(EA-IE2-1)//
TGAGCGGACTGCAAATCGCAACGCTACTTTCTCGAGAGCT (EA-IE2-2);
I-CRS, AATTCGAGCTCGCCCAGTGAACCGTCAGATCTAGAG (I-CRS-1)//
TCGACTCTAGATCTGACGGTTCACTIIICGAGCTCG (I-CRS-2);
G-CRS, AATTCGAGCTCGCCCAGTGAACCGTCAGATCTAGAG (G-CRS-1)//
TCGACTCTAGATCTGACGGTTCACTGGGCGAGCTCG (G-CRS-2);
PAL, AATTCGAGCTCGTTTAGTTAAACGTCTAGAG (PAL-1)//
TCGACTCTAGACGTTTAACTAAACGAGCTCG (PAL-2).

Preparation of IE86 protein

For procaryotic expression of the 86 kDa IE-2 (IE86) protein of HCMV, *Escherichia coli* strain M15 which contained the plasmid pREP4 (Diagen) was used. After transformation with the procaryotic IE86 expression plasmid pQE10IE2 containing a full-length IE-2 c-DNA (28), IE86 was expressed as a histidine-tagged protein and purified via metal-chelate affinity chromatography as described previously (28). The purified IE86 protein was used in gel retardation experiments, phenanthroline–copper footprinting and chemical interference analyses.

Gel retardation analysis

For gel retardation analyses the double-stranded oligonucleotides CRS, CRS-MUT, I-CRS, G-CRS and PAL were labeled by using polynucleotide kinase in the presence of γ-^{32}P-ATP. After purification of the probes via polyacrylamide gel electrophoresis, procaryotically expressed, purified IE86 protein (100ng) was mixed with 20 000 cpm (analytical reaction) or 100 000 cpm (preparative reaction) of the probe and the gel shift reaction was performed in buffer B containing 25 mM Tris (pH 8.0), 6.25 mM $MgCl_2$, 0.5 mM EDTA, 0.5 mM DTT, 0.01% Nonidet P-40, 9% glycerol and 1 μg of herring sperm DNA as an unspecific competitor. In competition experiments, 5 ng to 20 ng of the double-stranded CRS or PAL oligonucleotide were included in the reaction mix. Samples were electrophoresed as described (28).

Phenanthroline–copper protection analysis

For phenanthroline–copper (OP–CU) protection analysis the *Xho*I/*Bam*HI fragment of plasmid pHM69 containing the IE1/2 promoter region between nucleotides −38 and +6 was used as probe. This fragment was labeled either at the 3′ end of the *Bam*HI site by a filling in reaction with Klenow fragment (coding strand) or at the 5′ end of the *Bam*HI site by using polynucleotide kinase (non-coding strand). OP–CU protection analysis was performed as described by Kuwabara and Sigman (1987) (37) with minor modifications. After a gel retardation reaction and exposure of the wet gel to X-ray film for 1 h, the bands representing IE86/DNA-complexes and free DNA were excised from the gel and immersed in 600 μl of a solution containing 50 mM Tris/HCl (pH 8.0). Then, 60 μl of solution A containing 2.0 mM 1,10-phenanthroline (Sigma) and 0.45 mM $CuSO_4$ were added followed by 60 μl of solution B containing 58 mM β-mercaptopropionic acid (Sigma). After incubation for 10 min at room temperature the digestion reaction was quenched by the addition of 60 μl of 28 mM 2,9-dimethyl-OP (Sigma) and 1.4 ml of a solution containing 0.5 M ammonium acetate and 1 mM EDTA. The DNA was then eluted from the polyacrylamide gel onto a DEAE-membrane (Schleicher and Schüll) and electrophoresed under denaturing conditions. The gel was dried, autoradiographed and analysed by densitometry scanning.

Hydroxyl radical interference analysis

For hydroxyl radical interference (HRI) analysis the *Xho*I/*Bam*HI fragment of plasmid pHM69 was labeled as mentioned above. HR-interference reactions were carried out as described by Tullius and Dombroski (1986) (38) with modifications. About 1×10^5 cpm of end-labeled DNA in buffer B (see gel retardation analysis) and 1 μg of herring sperm DNA were mixed in a final volume of 50 μl. Then, 20 μl of freshly prepared cleavage solution (5 μl of 280 μM $Fe(NH_4)_2(SO_4)_2$, 5 μl of 520 μM EDTA, 5 μl 2% H_2O_2, and 5 μl 26.8 mM ascorbic acid) were added. After incubation for 20 min at room temperature, the reaction was stopped by the addition of 30 μl of 0.2 M thiourea and 1 μg of t-RNA. The DNA was precipitated, pelleted, dried under vacuum, and resuspended in buffer B. A preparative gel retardation experiment was performed and retarded and free DNA were eluted from the gel as described for OP–CU protection analysis. The resulting DNA fragments were separated via denaturing polyacrylamide gel electrophoresis. Afterwards, the gel was dried, autoradiographed and analysed by densitometry.

Methylation interference analysis

For methylation interference (MI) analysis the single-stranded oligonucleotides EAIE2-2, I-CRS-2, CRS-2, Pal-1 and Pal-2 were labeled by using polynucleotide kinase in the presence of γ-^{32}P-ATP. After heating to 95°C, the complementary strands EAIE2-1, I-CRS-1, CRS-1, PAL-2 and PAL-1 were added, respectively, and annealing was performed by gradual cooling of the samples to 4°C. The *Xho*I/*Bam*HI fragment of plasmid pHM69 was also used as a probe in MI-experiments (see OP–CU protection analysis). After purification via polyacrylamide gel electrophoresis, the end-labeled DNA-probes were modified by treatment with dimethylsulfate (DMS) as reported by Siebenlist and Gilbert (39). For preparative gel retardation

assays, modified DNA-probes (1×10^5 CPM) were incubated for 10 min in buffer B (see gel retardation analysis) together with 3 μg of herring sperm DNA as an unspecific competitor and 200 ng of purified IE86 protein in a final reaction volume of 50 μl. After native gel electrophoresis and exposure of the wet gel to X-ray film for 1 h, bound and free DNA were eluted from the gel matrix onto a DEAE membrane (Schleicher and Schuell) and cleaved at modified positions with piperidine. DNA fragments were separated using a denaturing polyacrylamide gel. Then, the gel was dried, autoradiographed and the autoradiogram was analysed by densitometry.

For OP−CU-, HRI- and MI-analysis, each protection or interference reaction was performed at least three times yielding a highly reproducible protection or interference pattern.

RESULTS

Delimitation of intimate contacts of IE86 with its target DNA by phenanthroline−copper protection analysis

The 86 kDa immediate early-2 (IE-2, IE86) protein of HCMV has previously been reported to negatively regulate its own expression via a short nucleotide sequence element, termed the *cis* repression signal (CRS). This nucleotide sequence is located between the TATA box and the cap site of the immediate early-1/2 (IE1/2) enhancer/promoter of HCMV and has been defined by mutational analysis to extend from nucleotides −14 to −1 relative to the transcriptional start site (see Figure 1) (25−27). In DNase I protection analysis a direct interaction of IE86 with the CRS element could be detected, thus demonstrating that IE86 is a DNA-binding protein (28). The protection pattern observed extended largely from nucleotides −18 to +10 on the coding strand and from nucleotides −25 to +3 on the non-coding strand relative to the IE-1/2 transcription start site, overlapping the TATA box of the IE1/2 promoter in part (see Figure 1). In further DNase I protection analysis using purified, procaryotically expressed IE86 protein and TATA box binding protein (TBP), we and others observed that both proteins are able to bind simultaneously at the IE-1/2 promoter in an immediate vicinity (data not shown) (36). Thus, a steric block of TBP-binding to the TATA box appeared not to be the major mechanism of IE86 mediated negative regulation. In addition, it suggested a smaller region of intimate contact of IE86 with its target DNA.

In order to define the protection pattern of IE86 with a higher precision, phenanthroline−copper (OP−CU) footprinting was performed (37). In the experiment of Figure 2, a DNA fragment comprising the IE1/2 promoter region between nucleotides −38 and +6 which was labeled either at the 3′ end of the coding strand (Fig. 2A) or the 5′ end of the non-coding strand (Fig. 2B) was incubated in a preparative gel retardation reaction with 200 ng of procaryotically expressed, purified IE86 protein. After electrophoresis and autoradiography, the bands representing the IE86/DNA complexes and the free DNA were excised from the gel and subjected to a light digestion treatment, so that each DNA-fragment was cleaved at no more than one position. DNA, eluted from the gel matrix, was electrophoresed on a sequencing gel and autoradiographed. The results were quantitated by densitometry scanning (Fig. 2C). Strong protections were visible on both the coding strand (Fig. 2A, lanes 3 and 4) and the non-coding strand (Fig. 2B, lanes 3 and 4). Nucleotides, protected on both strands, extended from −16 to −1 relative to the IE1/2 transcription start, exactly encompassing the CRS-element (Fig. 2C). In particular, the strongest protections relative to the IE1/2 transcription start extended from position −13 to −1 on the coding strand and from −14 to −4 on the non-coding strand. The flanking nucleotides were apparently less but significantly protected. Thus, the interacting region of 28 nucleotides as defined by DNase I footprinting assays was delimitated via OP−CU protection analysis to 16 nucleotides.

Hydroxyl radical interference analysis reveals the importance of single nucleoside positions for an interaction of IE86 with its autoregulatory target site

To further analyse the direct interactions between the IE86 protein and different bases of the DNA within the protected region, extending from nucleotides −16 to −1 relative to the IE1/2 transcription start, hydroxyl radical interference (HRI) analyses were performed. By hydroxyl radical cleavage individual nucleosides from a DNA fragment containing the IE1/2 promoter region between nucleotides −38 and +6 were removed. The DNA which was labeled either at the top or bottom strand was subjected to a light cleavage treatment, so that each DNA fragment was modified at no more than one position. IE86 protein

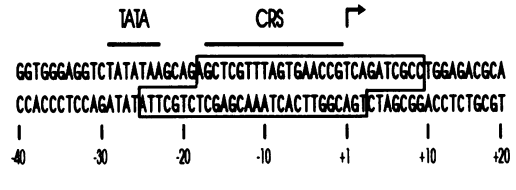

Figure 1. DNA sequence of the IE1/2 promoter region between nucleotides −40 to +22. Sequences protected from DNase I cleavage by the 86 kDa IE-2 protein are boxed. Filled bars indicate the location of the TATA-box and the CRS-element. The arrow shows the transcription start site of the IE1/2 transcription unit.

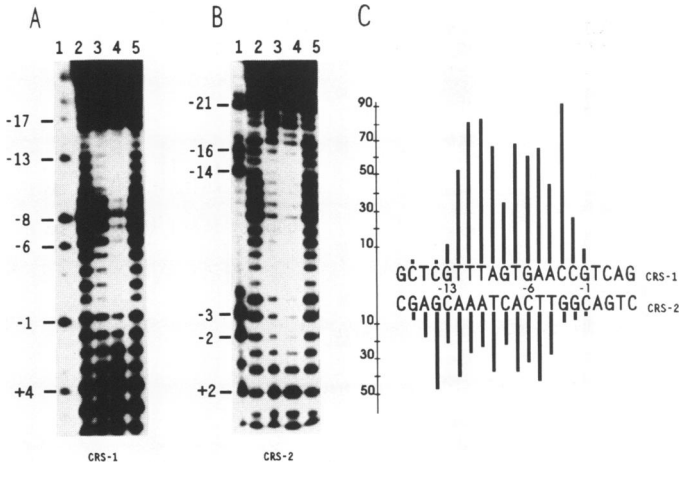

Figure 2. Phenanthroline−copper (OP−CU) protection analysis of the IE1/2 promoter region with the 86 kDa IE-2 protein. The *XhoI*/*Bam*HI fragment of plasmid pHM69 containing the IE1/2 promoter region was labeled either at the coding strand (**A**) or at the non-coding strand (**B**). The DNA probes were complexed with IE-2 protein, cleaved, isolated and displayed as mentioned in Materials and Methods. G-sequence reactions as a marker are shown in lanes 1. Lanes 2 and 5 show cleavage reactions without addition of IE-2 protein. Lanes 3 and 4 show cleavage reactions after addition of IE-2 protein. A summary of the results obtained in OP−CU protection analysis is shown in the histogram (**C**). The magnitude of the protection is shown by the size of the bar above and below the position of the DNA sequence.

was then incubated with the modified DNA, and the DNA fragments bound to IE-2 protein were separated from unbound fragments by native gel electrophoresis. DNA, eluted from the regions of the gel containing the bound and free fragments, was electrophoresed on a sequencing gel. The results are shown in Figure 3.

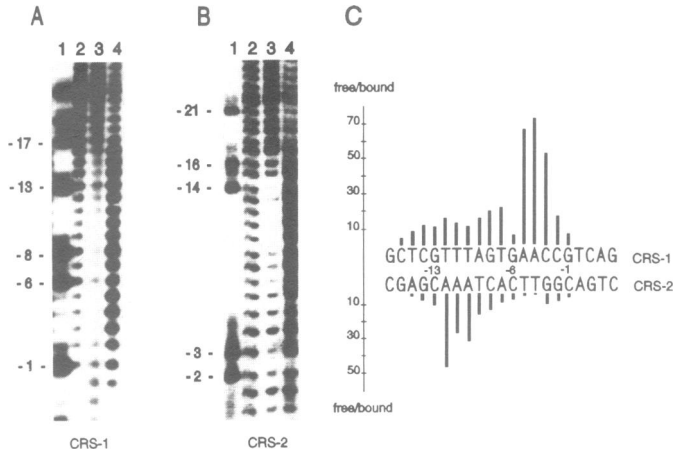

Figure 3. Hydroxyl radical interference analysis of the IE1/2 promoter region with the 86 kDa IE-2 protein. The *Xho*I/*Bam*HI fragment of plasmid pHM69 containing the IE1/2 promoter region was labeled at the coding strand (**A**) or at the non-coding strand (**B**). The DNA probes were modified, complexed with IE-2 protein and analysed as described in Materials and Methods. G-sequence reactions as a marker are shown in lanes 1. Input DNA without addition of IE-2 protein is shown in lanes 2. Lanes 3 and 4 represent the bound and free DNA, respectively, which was isolated following native gel electrophoresis. The data in (A) and (B) were analysed by densitometry and are summarized in the histogram (**C**). The ratio of the intensities of the bands observed in the bound and free DNA fractions was compared for each position of the CRS-element.

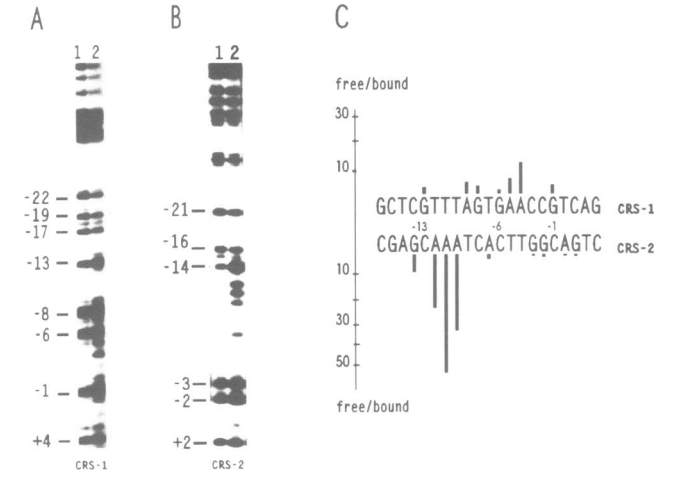

Figure 4. Methylation interference (MI) analysis of the IE1/2 promoter region with the 86 kDa IE-2 protein. The *Xho*I/*Bam*HI fragment of plasmid pHM69 containing the IE1/2 promoter region was labeled at the coding (**A**) or at the non-coding strand (**B**). The MI-pattern of the bound and free DNA isolated after a gel retardation assay is shown in lanes 1 and 2, respectively. DNAs were modified by DMS treatment, complexed with IE-2 protein, isolated, cleaved and displayed as described in Materials and Methods. The data in (A) and (B) were analysed by densitometry and are summarized in the histogram (**C**). The ratio of the intensities of the bands observed in the bound and free DNA fractions was compared for each position of the CRS-element.

Densitometry of the autoradiograms of Figures 3A (coding strand) and 3B (non coding strand) revealed the relative importance of individual nucleosides to the formation or maintenance of the IE86/CRS complex. These data are summarized in the histogram of Figure 3C. It is shown that the bases at different positions within the CRS-element have different relative importance. On the coding strand, the removal of the A's at positions −5 and −4 and the release of the C at position −3 strongly interfered with stable binding of IE86 protein to the DNA probe. In contrast, the bases at positions −16 to −6 and at −2 to −1 contributed to binding but were apparently less important. On the non-coding strand the A's at positions −12, −11, and −10 were the most important bases involved in IE86 binding relative to the flanking bases from positions −15 to −13 and from −9 to −1. Thus, the primary determinants of IE86 protein binding to the CRS-sequence appear to be contained in the 3′ direction of both the top and the bottom strand. Loci of maximal interference are shifted towards the 3′ boundaries. This HRI interference pattern of IE86 at the CRS-element resembles those of minor groove binding proteins (40,41).

Methylation interference analysis suggests minor groove contacts of IE86

To test the hypothesis of whether IE86 binds to the CRS-element via minor groove contacts, methylation interference (MI) analyses were performed. The primary advantage of this method is that information is obtained about whether the protein contacts the minor or major groove of DNA. This distinction is possible because adenines become methylated at the N-3 position within the minor groove and guanines at the N-7 position in the major groove. The DNA fragment containing the CRS-element, labeled as mentioned above, was partially modified by dimethyl sulfate (DMS) and incubated with the procaryotically expressed, purified IE86 protein. After gel retardation analysis, bound and free probes were eluted from the gel, treated to cleave the phosphodiester backbone at the position of each methylated base,

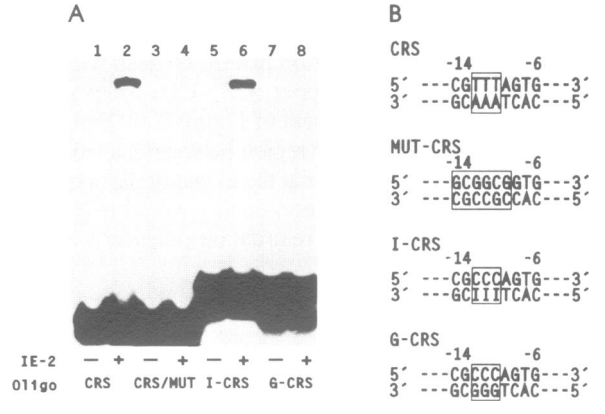

Figure 5. Gel retardation analysis with the inosine-substitued CRS-element (I-CRS) and the 86 kDa IE-2 protein. The nucleotide sequences of the oligonucleotides used in this assay are shown in (**B**). Identical amounts of IE-2 protein [(**A**), lanes 2, 4, 6 and 8] or buffer [(A), lanes 1, 3, 5 and 7] were incubated with the radiolabeled CRS (lanes 1 and 2), MUT-CRS (lanes 3 and 4), I-CRS (lanes 5 and 6) and G-CRS oligonucleotide (lanes 7 and 8).

and analysed by denaturing polyacrylamide gel electrophoresis. Figures 4A (coding strand) and 4B (non-coding strand) show representative results of the MI assays, and a quantitation of the results based on densitometry scanning is shown in Figure 4C. We found that methylation of several A's and G's within the CRS-element interfered with IE86 protein binding. In particular, methylation of any of the A's from −12 to −10 on the non-coding strand (Fig. 4A) and the A's at −5 and −4 on the coding strand (Fig. 4B) strongly interfered with IE86 binding, confirming the results obtained in HRI analysis. Methylation of the A at −9 and the G's at −13, −8, −6 and −1 on the coding strand together with the A's at −7 and +1 and the G's at −3, −2 and +2 on the non-coding strand was only mildly deleterious. The G at −14 on the non-coding strand showed relatively strong interference after methylation, indicating that the major groove is also contacted. The observed strong interference of methylated adenines on both strands of the DNA suggested that IE86 binds to the CRS-element via the minor groove.

IE86 protein binds specifically to an inosine/cytosine substituted CRS-element

To further confirm that IE86 protein binds DNA via minor groove contacts we decided to synthesize a double-stranded oligonucleotide containing an artificial CRS-element. This oligonucleotide, termed I-CRS, consisted of the CRS wild-type sequence except that the adenines at positions −12 to −10 of the non-coding strand were replaced by inosines (see Figure 5B). In terms of hydrogen bonding substituents on the purine and pyrimidine rings, TIIIC looks like TAAAC in the minor groove and like TGGGC in the major groove. If minor groove contacts were important for an interaction of IE86, the I-CRS oligonucleotide should be bound as well. To test this, we used both the I-CRS and the wild-type CRS oligonucleotide in a gel shift assay, comparing the retardation mediated by IE86. As control the G-CRS oligonucleotide was used (see Figure 5B). This oligonucleotide contained G's at positions −12 to −10 which should result in an altered minor groove morphology as well. In addition, a mutated form of the CRS oligonucleotide (Mut-CRS) was labeled which has previously been used as a negative control for IE86 DNA-binding (28). In the experiment of Figure 5 identical amounts of purified IE86 protein were incubated with the CRS-, mutated CRS- (CRS/MUT), I-CRS- and G-CRS oligonucleotide. IE86 protein retarded both the I-CRS- and the CRS-oligonucleotide with the same efficiency (Figure 5A, lanes 2 and 6). No retardation was observed with the MUT-CRS- and the G-CRS oligonucleotide (Figure 5A, lanes 4 and 8). Although these results established the minor groove as the primary determinant of IE86 binding to the CRS-element, we could not rule out participation of the major groove, for example, through recognition of the purine N-7 of inosines.

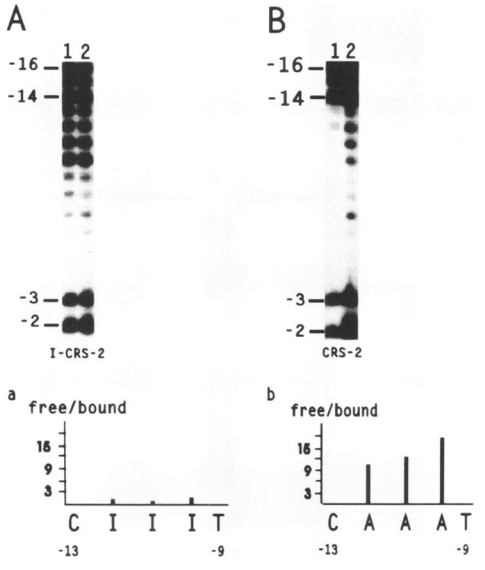

Figure 6. Methylation interference (MI) analysis of the I-CRS and the CRS-element with the 86 kDa IE-2 protein. Labeled I-CRS (**A**) and CRS oligonucleotides (**B**) were treated with DMS, complexed with IE-2 protein, isolated after a gel retardation reaction, cleaved, separated on a 12.5% sequencing gel and analysed by densitometry scanning. The MI-pattern of the bound and the free DNA of the I-CRS (A) and the CRS (B) element is shown in lanes 1 and 2, respectively. The results of (A) and (B) are summarized in the histograms a and b, respectively. The ratio of the intensities of the bands observed in the bound and free DNA fractions was compared for each position of the I-CRS and the CRS oligonucleotide.

Methylation in the major groove did not interfere with IE86 binding

The fact that IE86 protein bound strongly to the I-CRS oligonucleotide allowed us to determine directly the importance of major groove contacts. Inosine, like guanine, is methylated by DMS at the N-7 position within the major groove (42). Thus, if major groove contacts were not important for the interaction of IE86 with the I-CRS oligonucleotide, methylation at the inosines should not interfere with binding. In the experiment of Figure 6 the results of methylation interference analysis using the non-coding strand of both the I-CRS (Fig. 6A) and the CRS oligonucleotide (Fig. 6B) together with the quantitation based on densitometry scanning (Fig. 6C) are shown. Methylation of the three I's at positions −12 to −10 in the major groove caused a very weak interference, providing further confirmation that IE86 interacts with the I-CRS oligonucleotide primarily via minor groove contacts. In contrast, methylation at any of the three adenines at the same positions within the wild-type CRS-sequence showed strong interference (see Figures 5a and 5b). Thus, the methylation interference pattern observed for the I-CRS oligonucleotide is consistent with a model in which functional groups within the major groove of positions −12 to −10 do not contribute substantially to the binding of IE86 protein to the CRS-element.

Binding of IE86 within an early promoter of HCMV is also mediated via minor groove contacts

We demonstrated recently, that a direct interaction of IE86 with DNA is not confined to the CRS-element of the IE1/2 enhancer/promoter, but occurs also within an early promoter of HCMV (20). This promoter drives expression of the UL112/UL113 region of HCMV and is strongly transactivated upon cotransfection with IE86 expression vectors. Remarkably, the identified binding sites which were located within the region of −290 to −120 of the UL112 promoter, had a rather limited sequence similarity when compared to the CRS sequence. To test whether minor groove contacts are involved in the interaction of IE86 with one of these elements of the UL112 promoter, a double-standed oligonucleotide termed EAIE2 was synthesized. It corresponded in sequence to the IE86 binding site located between nucleotides −151 and −126 next to the TATA box of

the UL112 promoter. This oligonucleotide which was labeled at the 5′ end of the non-coding strand was used in methylation interference analysis. The result of this experiment is shown in Figure 7. Strong interference was observed after methylation of A's and G's. In particular, methylation of the three A's at positions −140 to −138 resulted in a strong interference. The flanking G at position −136 was moderately important whereas the G at −143 appeared to be more important for binding of IE86. Thus, the MI pattern of the EAIE2-2 oligonucleotide is in agreement with that observed for the CRS-element. The strong interference observed after methylation of the three adenines within both the CRS-element and the EAIE2 oligonucleotide confirmed that minor groove contacts are a major determinant of the interaction of IE86 with DNA.

IE86 binds to a palindromic CRS-element via symmetric minor groove contacts

The nucleotide sequence of the CRS-element is partially palindromic (see Figure 8B). To address the question as to whether the IE86 protein could bind to a palindromic CRS-element, a double-stranded oligonucleotide, termed PAL, was synthesized (see Figure 8B). This PAL oligonucleotide consisted of the CRS wild-type sequence with the exception that the second half-site GAACCG was replaced by TAAACG. To test whether IE86 could bind to the PAL-sequence, we used both the CRS and the PAL oligonucleotide in a gel retardation assay (Figure 8A). The IE86 protein retarded both oligonucleotides to the same position in the gel (Figure 8A, lanes 2 and 4). However, compared to the retardation observed with the CRS-probe, the retardation of the PAL oligonucleotide was three times more efficient, indicating a higher affinity for IE86. This could be confirmed by performing competition experiments: the PAL oligonucleotide competed much more efficiently for binding of IE86 to the CRS probe when compared to the wild-type CRS sequence (Figure 9C). This result shows that IE86 protein is able to bind to a palindromic CRS-sequence and indicates that contacts within both half sites determine the affinity of binding.

A methylation interference analysis was then performed in order to reveal the contacted nucleotide positions within the PAL oligonucleotide. Figures 9A and 9B, show the experimental results of the MI analysis of the coding (PAL-1) and non-coding (PAL-2) strand of the PAL oligonucleotide, respectively. A quantitation of the results based on densitometry scanning is shown in Figure 8C. Methylation of every adenine residue on both strands of the PAL oligonucleotide (at positions −7 and −5 to −3 on the PAL-1 strand and −12 to −10 and −7 to −6 on the PAL-2 strand) strongly interfered with IE86 binding. The G's at positions −13, −8 and −1 on PAL-1 and at positions −14 and −2 on PAL-2 showed also interference after DMS treatment. These data indicate that IE86 protein can bind to a palindromic sequence element via symmetric interactions within the minor groove of DNA.

DISCUSSION

The 86 kDa IE2 (IE86) protein of HCMV is a multifunctional polypeptide that can mediate both transactivation and repression of gene transcription (9,15,16,24,43). Repression is specific for

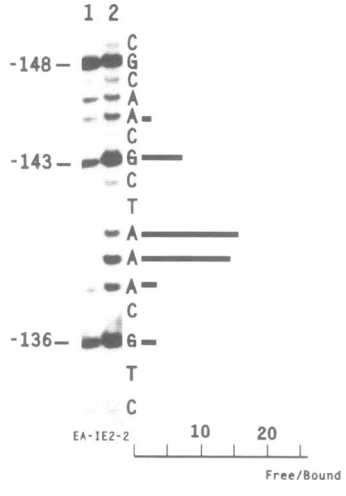

Figure 7. Methylation interference analysis with the EA-IE2 oligonucleotide and the 86 kDa IE-2 protein. The EA-IE2 oligonucleotide was radiolabeled at the EA-IE2-2 strand and treated with DMS. After a gel retardation reaction with IE-2 protein, bound and free DNA was isolated, cleaved and displayed as described in Materials and Methods. The MI-pattern of the bound and free DNA is shown in lanes 1 and 2, respectively. The nucleotide sequence of the EA-IE2-2 oligonucleotide and the results of densitometric analysis are shown as indicated.

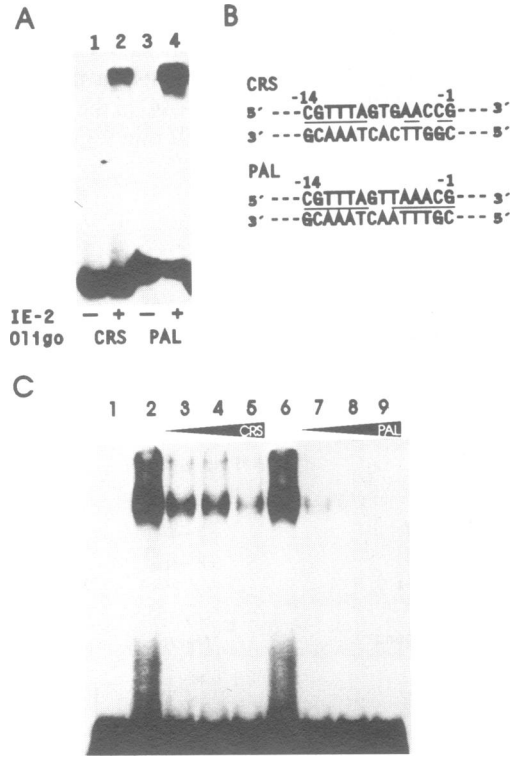

Figure 8. Gel retardation analysis with the palindromic CRS-element and the 86 kDa IE-2 protein. The nucleotide sequences of the CRS and the palindromic CRS-element (PAL) are shown in (**B**). Identical amounts of IE-2 protein[(**A**), lanes 2 and 4] or buffer [(A), lanes 1 and 3] were incubated with the radiolabeled CRS (lanes 1 and 2) or the PAL oligonucleotide (lanes 3 and 4). In (**C**) a competition analysis is shown. The CRS oligonucleotide was used as a probe in gel retardation analysis either without protein [(C), lane 1] or together with identical amounts of IE-2 protein [(C), lanes 2 to 9]. Either the unlabeled CRS oligonucleotide [5 ng, 10 ng or 20 ng, respectively; see (C), lanes 3 to 5] or the unlabeled PAL oligonucleotide [5 ng, 10 ng or 20 ng, respectively; see (C), lanes 7 to 9] were used for competition analysis.

the major IE enhancer/promoter within the genome of HCMV and requires the presence of a sequence element, termed *cis* repression signal (CRS) that is located between the TATA-box and the transcription start site (25–27). In this study we demonstrate that IE86 interacts with the CRS sequence primarily through contacts within the minor groove of the helix. Moreover, this appears to be a general characteristic of the IE86/DNA interaction as contacts with the minor groove occur also at an IE86 binding site located within an early promoter of HCMV. Thus, IE86 is a member of a small group of DNA binding proteins that interact with A–T rich sequences within the minor groove. This group includes polypeptides such as the TATA-box binding protein TBP (31–34), the HMG domain protein LEF-1 (35) and the bacterial protein IHF (44). Minor groove interaction is an unusual property as the minor groove has been thought to contain insufficient information to define a sequence specific recognition site for a protein (45). Indeed, most DNA-binding proteins recognize the hydrogen bonding surface within the major groove. A detailed description of the interaction of IE86 with its target sites may therefore contribute information about the determinants of protein/DNA interaction. In addition, our results have important implications concerning the function of IE86 in regulation of HCMV gene transcription.

Three lines of evidence support our conclusion that IE86 binds within the minor groove of DNA. Firstly, we have used hydroxyl radical interference analysis to identify the nucleosides within the CRS-element that contribute directly to the formation and maintenance of the IE86/CRS complex. The results obtained in this assay and in phenanthroline copper protection experiments showed that the bases of maximal protection are shifted to the 3′ boundaries; this protection pattern is characteristic of protein/DNA interactions in the minor groove (40,41). Secondly, we found that methylation of the A residues on both DNA strands within the CRS-element interfered strongly with IE86 binding. Methylation of A alters the minor groove surface both by eliminating a potential hydrogen bond acceptor and by introducing a novel group that might interfere sterically. However, modified

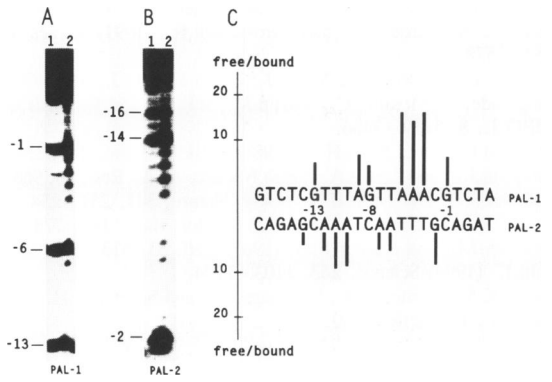

Figure 9. Methylation interference (MI) analysis of the palindromic CRS-element (PAL) with the 86 kDa IE-2 protein. PAL oligonucleotide labeled on both the PAL-1 and the PAL-2 strand was treated with DMS, complexed with IE-2 protein, isolated, cleaved, separated on a 12,5% sequencing gel and analysed by densitometry scanning. The MI-pattern of the bound (lane 1) and the free DNA (lane 2) of the PAL-1 and the PAL-2 strand is shown under **A** and **B**, respectively. The data of densitometry scanning of both strands of the PAL element are summarized in the histogram (**C**). The ratio of the intensities of the bands observed in the bound and free fractions was compared for each position within the PAL oligonucleotide.

bases might also interfere with protein binding by altering the proclivity of the DNA to adopt a necessary structure (46,47). For this reason we decided to employ the approach of Starr and Hawley (1991) (31). We substituted the most important A–T residues within the CRS-sequence for I-C bases in order to alter the recognition potential of the major groove while leaving the minor groove unchanged. As demonstrated by gel retardation analysis IE86 binds to this I-substituted CRS element at a level similar to that of the wild-type CRS-sequence. This revealed that major groove functional groups were relatively unimportant for an interaction of IE86 with DNA. This conclusion was strengthened further by the observation that methylation of the I's in the major groove caused only a small decrease in IE86 binding. However, few major groove contacts also appear to participate in specific DNA-binding of IE86, because methylation of the N7 position of one or two G residues flanking the A–T contact sites interfere with binding. In this respect IE86 resembles TBP and LEF-1 which recognize their DNA binding sites primarily through the minor groove together with a few major groove contacts in the flanking positions (31,32,35).

One major characteristic of minor groove binding proteins is their ability to interact with a wide variety of DNA sequences. This has been reported for TBP (48), for HMG-box proteins (49) and for IHF, which is a sequence-specific DNA-binding protein but binds to nonspecific DNA under certain conditions (50). A similar variation in sequence-specific recognition of DNA is observed with IE86. Firstly, as described for TBP, IE86 binds to synthetic polynucleotides like poly (dI-dC)·poly (dI-dC) (28,30). This could now be explained by an interaction of IE86 with this unspecific competitor DNA via minor groove contacts. Secondly, IE86 has three target sites within the UL112 promoter of HCMV which have little sequence homology when compared to the CRS-element, indicating that IE86 can interact with a broad spectrum of DNA sequences (20). However, substitution of the most important A–T residues within the CRS-element for G–C, which alters the minor groove surface of DNA, is deleterious for IE86 binding. This again resembles the DNA binding properties of TBP and LEF-1 (35,48,51). The only nucleotides that are conserved between all IE86 binding sites identified so far are two copies of the dinucleotide C–G or G–C separated by ten nucleotides. In addition to the A–T-stretches flanking the C–G or G–C dinucleotides internally, this spacing is probably a further determinant of the IE86/DNA interaction, because an insertion mutant of one nucleotide within the CRS-element abolished IE86 binding (30).

We have shown that IE86 interacts with a palindromic CRS-element. Binding appears to be more efficient when compared to the wild-type CRS-sequence. This result makes dimerisation of IE86 while binding to DNA possible. It is worth to note that increasing the amount of purified IE86 protein in gel retardation reactions results in the appearence of slower migrating IE86/DNA complexes (unpublished). This indicates that IE86 is able to multimerize when bound to DNA. In addition, it is known that IE86 can interact with itself in solution (23,30). However we and others could not successfully demonstrate in gel retardation analysis by using truncated forms of the protein that the lower migrating complex represents an IE86 dimer (unpublished) (30). In part, this may be due to the fact that the domain of IE86 which is minimally required for DNA binding still comprises more than 200 amino acids (30). A comparison of this amino acid sequence with well-characterized minor groove binding motifs such as the HMG domain did not reveal any substantial homologies.

As a further important characteristic, minor groove binding proteins have been shown to bend DNA heavily. This appears to be of functional relevance; for example, IHF promotes the formation of a compact structure required for integration of phage lambda into the host chromosome (52–54). In the case of TBP, the DNA at the TATA box is strongly bent (100°) in a direction opposite to that preferred in nucleosomal DNA (33,34,55), suggesting a simple mechanism for the observed competition between histones and TBP. This has been described as an important rate-determining step in transcription (56,57). In this respect, IE86 has been reported to counteract histone H1-dependent repression of the UL112 promoter of HCMV (58) which, as detected recently (20), contains binding sites for IE86. Remarkably, as IE86 and TBP are able to bind simultaneously at the major IE enhancer/promoter of HCMV (36), IE86 induced bending at the CRS-element immediately downstream of the TATA box could disturb the TBP induced bending pattern and thus constitute a potential mechanism for repression of transcription from this promoter.

Taken together we have demonstrated that the 86 kDa IE-2 protein of HCMV binds to DNA primarily via minor groove contacts. This finding extends explanations of both IE86 dependent transactivation as well as repression of gene transcription.

ACKNOWLEDGEMENTS

We thank B.Fleckenstein for critical reading of the manuscript and continuous support. This work was supported by the Deutsche Forschungsgemeinschaft (Forschergruppe Transkriptionssignale und Regulatorproteine) and the Sophie-Wallner-Fonds.

REFERENCES

1. Reichert,C.M., O'Leary,T.J., Levens,D.L., Simrell,C.R. and Macher,A.M. (1983) Am. J. Pathol., **112**, 357–382.
2. Chee,M.S., Bankier,A.T., Beck,S., Bohni,R., Brown,C.M., Cerny,R., Horsnell,T., Hutchison,C.A., Kouzarides,T., Martignetti,J.A. and *et al*, (1990) Curr. Top. Microbiol. Immunol., **154**, 125–169.
3. Wathen,M.W., Thomsen,D.R. and Stinski,M.F. (1981) J. Virol., **38**, 446–459.
4. Wathen,M.W. and Stinski,M.F. (1982) J. Virol., **41**, 462–477.
5. McDonough,S.H. and Spector,D.H. (1983) Virology, **125**, 31–46.
6. Colberg-Poley,A.M., Santomenna,L.D., Harlow,P.P., Benfield,P.A. and Tenney,D.J. (1992) J. Virol., **66**, 95–105.
7. Stasiak,P.C. and Mocarski,E.S. (1992) J. Virol., **66**, 1050–1058.
8. Tevethia,M.J., Spector,D.J., Leisure,K.M. and Stinski,M.F. (1987) Virology, **161**, 276–285.
9. Hermiston,T.W., Malone,C.L., Witte,P.R. and Stinski,M.F. (1987) J. Virol., **61**, 3214–3221.
10. Pizzorno,M.C., Mullen,M.A., Chang,Y.N. and Hayward,G.S. (1991) J. Virol., **65**, 3839–3852.
11. Stenberg,R.M., Depto,A.S., Fortney,J. and Nelson,J.A. (1989) J. Virol., **63**, 2699–2708.
12. Boshart,M., Weber,F., Jahn,G., Dorsch-Hasler,K., Fleckenstein,B. and Schaffner,W. (1985) Cell, **41**, 521–530.
13. Thomsen,D.R., Stenberg,R.M., Goins,W.F. and Stinski,M.F. (1984) Proc. Natl. Acad. Sci. USA, **81**, 659–663.
14. Stenberg,R.M., Witte,P.R. and Stinski,M.F. (1985) J. Virol., **56**, 665–675.
15. Malone,C.L., Vesole,D.H. and Stinski,M.F. (1990) J. Virol., **64**, 1498–1506.
16. Pizzorno,M.C., O'Hare,P., Sha,L., Lafemina,R.L. and Hayward,G.S. (1988) J. Virol., **62**, 1167–1179.
17. Hagemeier,C., Walker,S.M., Sissons,P.J. and Sinclair,J.H. (1992) J. Gen. Virol., **73**, 2385–2393.
18. Hagemeier,C., Walker,S., Caswell,R., Kouzarides,T. and Sinclair,J. (1992) J. Virol., **66**, 4452–4456.
19. Caswell,R., Hagemeier,C., Chiou,C.-J., Hayward,G.S., Kouzarides,T. and Sinclair,J.H. (1993) J. Gen. Virol., **74**.
20. Arlt,H., Lang,D., Gebert,S. and Stamminger,T. (1994) J. Virol., **68**, 4117–4125.
21. Staprans,S.I., Rabert,D.K. and Spector,D.H. (1988) J. Virol., **62**, 3463–3473.
22. Chang,C.P., Malone,C.L. and Stinski,M.F. (1989) J. Virol., **63**, 281–290.
23. Furnari,B.A., Poma,E., Kowalik,T.F., Huong,S.M. and Huang,E.S. (1993) J. Virol., **67**, 4981–4991.
24. Hermiston,T.W., Malone,C.L. and Stinski,M.F. (1990) J. Virol., **64**, 3532–3536.
25. Liu,B., Hermiston,T.W. and Stinski,M.F. (1991) J. Virol., **65**, 897–903.
26. Cherrington,J.M., Khoury,E.L. and Mocarski,E.S. (1991) J. Virol., **65**, 887–896.
27. Pizzorno,M.C. and Hayward,G.S. (1990) J. Virol., **64**, 6154–6165.
28. Lang,D. and Stamminger,T. (1993) J. Virol., **67**, 323–331.
29. Macias,M.P. and Stinski,M.F. (1993) Proc. Natl. Acad. Sci. USA, **90**, 707–711.
30. Chiou,C.J., Zong,J., Waheed,I. and Hayward,G.S. (1993) J. Virol., **67**, 6201–6214.
31. Starr,D.B. and Hawley,D.K. (1991) Cell, **67**, 1231–1240.
32. Lee,D.K., Horikoshi,M. and Roeder,R.G. (1991) Cell, **67**, 1241–1250.
33. Kim,J.L., Nikolov,D.B. and Burley,S.K. (1993) Nature, **365**, 520–527.
34. Kim,Y., Geiger,J.H., Hahn,S. and Sigler,P.B. (1993) Nature, **365**, 512–520.
35. Giese,K., Cox,J. and Grosschedl,R. (1992) Cell, **69**, 185–195.
36. Jupp,R., Hoffmann,S., Depto,A., Stenberg,R.M., Ghazal,P. and Nelson,J.A. (1993) J. Virol., **67**, 5595–5604.
37. Kuwabara,M.D. and Sigman,D.S. (1987) Biochemistry, **26**, 7234–7238.
38. Tullius,T.D. and Dombroski,B.A. (1986) Proc. Natl. Acad. Sci. USA, **83**, 5469–5473.
39. Siebenlist,U. and Gilbert,W. (1980) Proc. Natl. Acad. Sci. USA, **77**, 122–126.
40. Van Dyke,M.W., Hertzberg,R.P. and Dervan,P.B. (1982) Proc. Natl. Acad. Sci. USA, **79**, 5470–5474.
41. Churchill,M.E. and Travers,A.A. (1991) Trends. Biochem. Sci., **16**, 92–97.
42. Scheit,K.-H. and Holy,A. (1967) Biochim. Biophys. Acta, **149**, 344–354.
43. Stenberg,R.M., Fortney,J., Barlow,S.W., Magrane,B.P., Nelson,J.A. and Ghazal,P. (1990) J. Virol., **64**, 1556–1565.
44. Yang,C.C. and Nash,H.A. (1989) Cell, **57**, 869–880.
45. Seeman,N.C., Rosenberg,J.M. and Rich,A. (1976) Proc. Natl. Acad. Sci. USA, **73**, 804–808.
46. McClarin,J.A., Frederick,C.A., Wang,B.C., Greene,P., Boyer,H.W., Grable,J. and Rosenberg,J.M. (1986) Science, **234**, 1526–1541.
47. Lu,A.L., Jack,W.E. and Modrich,P. (1981) J. Biol. Chem., **256**, 13200–13206.
48. Wobbe,C.R. and Struhl,K. (1990) Mol. Cell Biol., **10**, 3859–3867.
49. Giese,K., Amsterdam,A. and Grosschedl,R. (1991) Genes Dev., **5**, 2567–2578.
50. Bonnefoy,E. and Rouviere Yaniv,J. (1991) EMBO J., **10**, 687–696.
51. Giangrande,A., Mettling,C., Martin,M., Ruiz,C. and Richards,G. (1989) EMBO J., **8**, 3459–3466.
52. Robertson,C.A. and Nash,H.A. (1988) J. Biol. Chem., **263**, 3554–3557.
53. Thompson,J.F. and Landy,A. (1988) Nucleic. Acids. Res., **16**, 9687–9705.
54. Goodman,S.D. and Nash,H.A. (1989) Nature, **341**, 251–254.
55. Drew,H.R. and Travers,A.A. (1985) J. Mol. Biol., **186**, 773–790.
56. Workman,J.L. and Roeder,R.G. (1987) Cell, **51**, 613–622.
57. Struhl,K. (1994) Science, **263**, 1103–1104.
58. Klucher,K.M., Sommer,M., Kadonaga,J.T. and Spector,D.H. (1993) Mol. Cell Biol., **13**, 1238–1250.