Domains of the Epstein-Barr virus (EBV) transcription factor R required for dimerization, DNA binding and activation

Evelyne Manet, Agnès Rigolet, Henri Gruffat, Jean-François Giot and Alain Sergeant*
Laboratoire de Virologie Moléculaire, Ecole Normale Supérieure de Lyon, UMR 49 CNRS-ENS; 46 Allée d'Italie, 69364 Lyon Cedex 07, France

Received February 21, 1991; Revised and Accepted April 17, 1991

ABSTRACT

In cells latently infected with EBV, the switch from latency to a productive infection is linked to the expression of two transcriptional activators, the upstream element factor EB1 and the enhancer factor R. R activates by interacting directly with specific DNA sequences called RREs (R Responsive Elements). Each binding site covers about 18 bp, where R simultaneously contacts two core sequences separated by 5 to 7 bp (1). Here we show that R binds in vitro as a homodimer to an RRE, and that stable homodimers can also form in solution in the absence of DNA. By functional analysis of deletion and insertion mutants of R, we have localized the DNA binding region within the 280 N-terminal amino acids and the dimerization region within the 232 N-terminal amino acids. As no obvious homologies were detected with other known DNA binding or dimerization motifs, R could contain novel protein structures mediating these functions. The transcriptional activation domain has been located in the C-terminal half of the protein. This domain contains two regions with structures already identified in other transcription factors: one region is rich in proline, the other rich in acidic residues.

INTRODUCTION

The human herpes virus EBV (Epstein-Barr Virus) infects and immortalizes peripheral B lymphocytes, resulting in the establishment of a latent infection. In such latently infected B cells, the entire EBV genome is maintained largely as a plasmid, and its expression is reduced to a few genes: those encoding two small RNAs (EBERS) (2), the six Epstein-Barr Nuclear Antigens (EBNA-1, -2, -3A, -3B, 3C and LP; for a review see ref 3), the BHRF1 encoded protein (4; 5), the latent membrane protein (LMP) (6) and the terminal membrane proteins (TP1 and TP2), whose coding sequences are created by joining the ends of the linear virus (7; 8).

Since EBV is a lytic virus, the maintenance of latency is a condition for persistence of immortalization. However, in particular cell lines, between 0.5% and 5% of the cells spontaneously produce virus. Moreover, various chemical agents, including the tumour promoter 12-O-tetradecanoyl-phorbol-13-acetate (TPA), can cause the virus to switch from a latent to a lytic replication cycle (9). In both cases, the activation seems to be linked to the expression of two EBV-encoded transactivators: the BZLF1-encoded factor, EB1 (also called Z, Zta or ZEBRA) and the BRLF1-encoded factor, called R (or Rta)(10; 11; 12; 13; 14; 15; 16). EB1, which seems to have a key role in the induction of the lytic cycle, is a DNA binding protein that activates transcription from different responsive elements including AP-1 binding sites (17; 18; 19; 20; 21). EB1 does not seem to be a factor that can act at distances greater than 100 to 200 bp from the TATA box (20).

R, conversely, seems to be a factor that can act at distances over thousands of base pairs and several R targets (called RREs, for R Responsive Elements) have been identified (22; 23a; 23b; 24; 25; 26). One is part of the duplicated promoter DR/DL (RRE-DR) (27; 22), and overlaps with the enhancer of the EBV origins of replication active only during the lytic cycle (ORIlyt) (28). Recently, we have shown that in vitro R binds independently to two sites within the RRE-DR (the RRE-DR1 and the RRE-DR2). Each binding site covers about 18 bp, where R probably simultaneously contacts two core sequences separated by 5 to 7 bp (1). A third R-binding site (RRE-M) has been localized in the BMLF1 R-responsive enhancer (Gruffat et al, submitted for publication). This RRE-M has some homologies with RRE-DR1 and RRE-DR2, but these three binding sites are imperfectly conserved.

Here we report a detailed analysis of the functional domains of the 605 amino acid R protein. The region necessary for direct interaction with specific DNA sequences is localized in the N-terminal 280 amino acids. We show that R binds in vitro to specific DNA sequences as a homodimer and that the 232 N-terminal amino acids are necessary for R to form stable homodimers in solution. We have also localized regions of the protein responsible for the activation function. These regions are localized in the C-terminal half of the protein, and have no obvious role in the DNA binding capacity of R.

* To whom correspondence should be addressed

MATERIALS AND METHODS

Construction of insertion and deletion mutants of R

Plasmid pKSVR containing the R coding sequence (29) was cut with restriction enzymes Alu I, Hae III, Hinc II, Sma I and Pvu II in the presence of Ethidium Bromide so as to produce a maximum of single cut linear molecules, as described by Everett et al. (30). Linear DNA was isolated after agarose gel electrophoresis and ligated directly to 12 bp EcoR I linkers (Biolabs). The insertion location and the number of linkers inserted was determined by sequencing. When necessary, plasmids were extensively digested with EcoR I, religated and sequenced to ensure that only one linker was inserted.

Internal deletion mutants were generated by cutting different insertion mutants with both EcoR I and a second enzyme with a unique site in the vector, then religating the appropriate DNA fragments to generate in phase internal deletions. Out of frame deletion mutants were converted to in frame by recutting with EcoR I and religating, after blunting the extremities with the Klenow fragment of DNA-polymerase.

C-terminal deletion mutants were generated by ligating the following double stranded oligonucleotide

5'-AATTCTAGGTAGGTAGCTGCAGC-3', which contains stop codons in all three open reading frames, into the EcoR I sites present in the insertion mutants.

The d352-408 deletion mutant was made by site-directed mutagenesis using the MUTA-GENE M13 in vitro mutagenesis kit from Bio-Rad. All the mutants produced were sequenced before use.

Production of R and mutant proteins in vitro

The pSPTR construct was obtained by cloning a Dra I fragment containing the whole BRLF1 open reading frame into the Sma I site of pSPT19. Mutated R coding fragments were subcloned into plasmid pSPT19. Template DNAs for in vitro transcription were prepared by cutting in the pSPT19 polylinker, 3' from the R open reading frame, or with restriction enzymes (Hae II, Sac I, Sma I, Xho I) that cut within the coding region. Capped RNA was synthesized by the SP6 Polymerase, using standard conditions (Boehringer Mannheim).

The pT7βRd2-22 plasmid was generated by subcloning the R coding fragment truncated in 3', into plasmid pT7βSal/Stu (31) so that the AUG provided by the vector was in frame with the truncated BRLF1 open reading frame. For RNA synthesis, the template DNA was digested with EcoR I then transcribed by the T7 polymerase.

The pSPTRdm construct was generated in two steps. First a Nae I restriction site was introduced, by site-directed mutagenesis, after the last codon of the R protein. Second, a Sma I-Pst I fragment coming from the EB1 sequence was cloned into this modified pSPTR plasmid, between the newly created Nae I site and the unique Pst I site in the pSPT19 polylinker. Template DNA was digested with Pst I prior to synthesis with the SP6 polymerase.

pSPTEB1 was obtained by subcloning an EB1 cDNA into plasmid pSPT19.

In vitro transcribed RNA was used to initiate protein synthesis in messenger-dependent rabbit reticulocyte lysates (Promega) using either ^{14}C-Leucine or ^{35}S-Methionine.

Electrophoretic Mobility Shift Assays (EMSA)

1 to 5 μl of in vitro translation extract, was incubated with 1.5×10^5 cpm of ^{32}P-labeled probe containing the R binding site RRE-DR1 (1). Incubations were carried out in 0.5 mM MgCl2, 10 mM HEPES-KOH (pH 7.9), 0.5 mM DTT, 0.5 mM PMSF, 150 mM KCl, 10% glycerol, at 25°C for 30 minutes, in the presence of 1 μg of poly dI-dC non-specific competitor. The mixture was loaded onto a 4.5% polyacrylamide gel (crosslinked 29 to 1) with 0.2×TBE. The protein-DNA complexes were separated from the non-complexed DNA by migration at 10 V/cm and visualised by autoradiography.

Immunoprecipitations

10 μl of in vitro translated extract was incubated in 100 μl of buffer A (Tris 0.05 M, pH 7.2; NaCl 0.15 M; EDTA 1 mM; Nonidet P40 0.1%), with 1.5 μl of anti-EB1 antibody (32) for 1.5 hour at 4°C. 50 μl of protein A sepharose beads (Sigma) resuspended at 30 mg/ml was then added for a further hour incubation at 4°C. The beads were then washed three times in buffer A. The immunoprecipitated proteins were released by heating before loading onto SDS-PAGE.

Cell culture and transfections

HeLa cells were grown in DMEM (Gibco) supplemented with 10% (v/v) fetal calf serum. The plasmids used for transfection were prepared by the alkaline lysis method and purified through two CsCl gradients. HeLa cells were seeded at 10^6 cells per 100 mm Petri dish 8 h prior to transfection. Transfections were performed using the calcium precipitate method (33). Cells were mixed with the appropriate DNA(s), and the DNAs were in the same topological state as assayed by agarose gel electrophoresis. Usually 15 μg of DNA were used per 100 mm dish including pSV2β as internal control and pUC18 up to the 15 μg total.

RNA extraction and S1 nuclease mapping

The transfected cells were lysed using NP40 as described elsewhere (34). Nuclei were pelleted and RNA phenol extracted from the cytoplasmic fraction. 10 to 40 μg of total cytoplasmic RNA was hybridized overnight at 30°C in 50% formamide, 0.3M NaCl, 0.01M Tris-HCl pH 7.4 to 5'- ^{32}P-labelled single-stranded DNA probes. The hybrids were digested for 2 hours at 20°C with 5U of S1 nuclease per 10 μg of RNA. The size of the S1 protected DNA fragment was analysed on 8% (w/v) acrylamide/8.3 M urea gels. Quantification was made by cutting the specific S1-protected bands out of the gel and counting the radioactivity. The results were corrected as follows: (i) according to the efficiency of transfection as evaluated by counting the radioactivity present in S1-protected probes corresponding to specific SV40 early RNA expressed from plasmid pSV2β and (ii) according to the activities of the different constructions in the absence of R.

RESULTS

The DNA binding region of R is localized within the 280 N-terminal amino acids

We have already reported that in vitro R binds directly to specific DNA sequences, and that the DNA binding domain is contained within the 356 N-terminal amino acids (1). To refine the location of the DNA binding region, we produced a series of R mutants (Figure 1A). These mutated proteins were synthesized in vitro from RNA transcribed from plasmid pSPTR and its derivatives (see Materials and Methods), and assayed for their capacity to form specific complexes with a double-stranded oligonucleotide containing one of the R-binding sites identified in the ORI_{lyt}

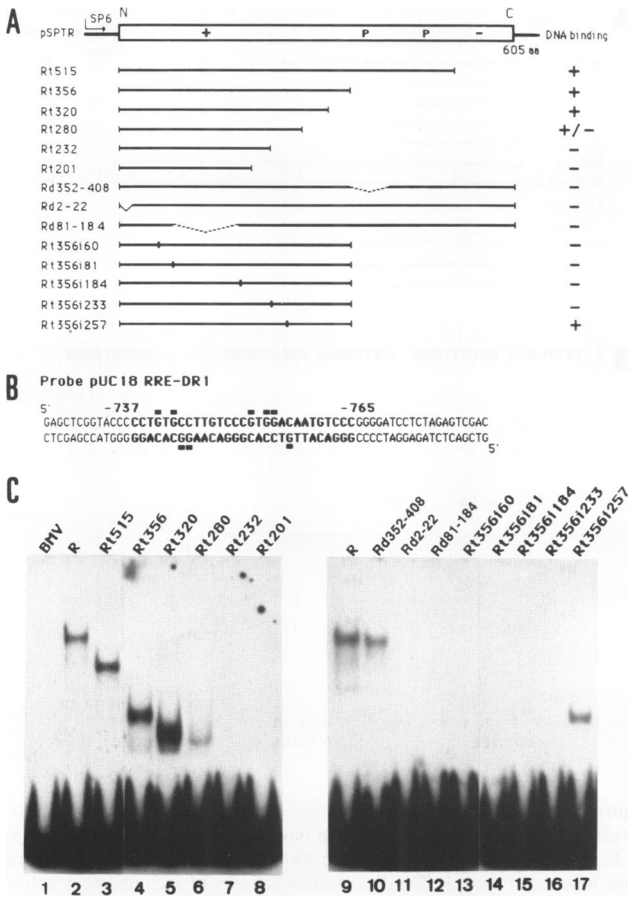

Figure 1. Mapping of the DNA binding domain. (A) Schematic representation of truncated and deleted R mutants that have been tested for binding. The plasmid pSPTR is shown schematically in the upper line. The coding region is depicted as a box and dominant characteristics of the amino acid sequence are indicated by the following symbols: + for a basic region, P for a region rich in proline residues, − for an acidic region. The names of the truncated mutants take the form RtX, where × represents the number of N-terminal codons remaining in the protein. Sequences between × and Y are missing in internal deletion mutants denoted as RdX-Y. Finally, proteins resulting from insertions in the Rt356 mutant are referred to as Rt356iX, where × indicates the site of the insertion. The DNA binding activity of each mutant is indicated (+ or −). +/− indicates only weak binding. (B) Sequence of the double stranded probe used for the EMSA. The sequence in bold type corresponds to enhancer sequences while the normal type indicates the surrounding sequences of pUC18. The coordinates −737 and −765 indicate the position of the sequence in the DR enhancer (start of transcription: +1). The G residues which have previously been shown to be contacted by the R protein are indicated by black squares (1). (C) Analysis by EMSA of the binding of the R mutants The proteins were translated in vitro and labelled with ^{14}C-Leucine. TCA precipitation of the proteins and subsequent radioactivity counting was carried out to ensure that the same amount of each protein was used for incubating with the ^{32}P-labeled RRE-DR1 probe.

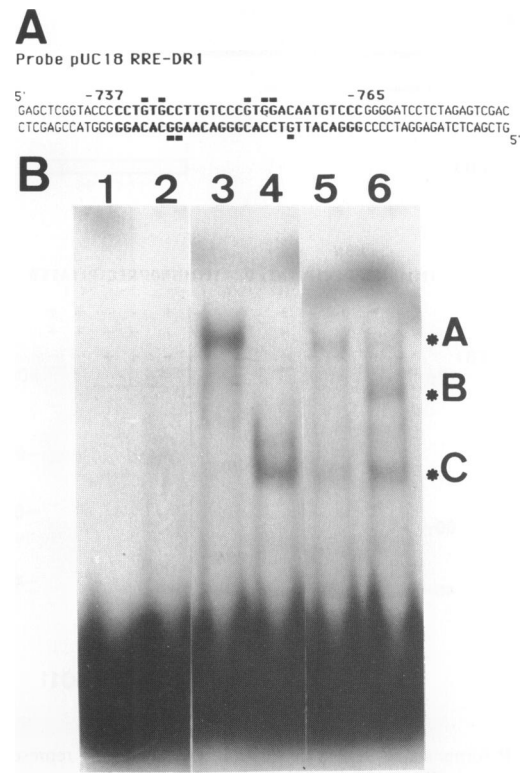

Figure 2. R binds DNA in vitro as a dimer. (A) Sequence of the RRE-DR1 probe used for the EMSA. (B) The wild-type R protein and the C-terminal truncated mutant Rt356 were translated in vitro in the presence of ^{14}C-Leucine, either together or separately, then mixed in the binding assay. The probe was incubated with no added protein (lane 1), Brome Mosaic Virus (BMV) proteins (lane 2), wild-type R proteins (lane 3), Rt356 mutant proteins (lane 4), R and Rt356 proteins synthesised separately and mixed before the incubation (lane 5), R and Rt356 proteins translated together (lane 6). R and Rt356 putative homodimers are indicated respectively by A and C, and the R/Rt356 heterodimers by B.

enhancer (RRE-DR1, Figure 1B) (1). The same amount of each protein was used in the assay, as estimated both by radioactive counting after precipitation of the proteins by TCA and by comparison of band intensity on an autoradiogramm after SDS-PAGE (not shown).

As shown in Figure 1C, progressive deletions in the carboxy-terminal of the R protein could be made until amino acid 320 without affecting the DNA binding activity (lanes 2 to 5). Further deletion of amino acids 320 to 280 decreased the amount of specific complexes formed with RRE-DR1 (lane 6), and no binding activity was detectable when deletions were extended beyond amino acid 280 (lanes 7 and 8). Deleting the first 20 N-terminal amino acids of the R protein also completely impaired its specific binding to DNA (lane 11).

Internal deletion in the C-terminal part of the R protein (mutants Rd352-408), did not affect its capacity to bind to DNA (lane 10), whereas the internal deletion in the N-terminal part (mutant Rd81-184) completely impaired formation of protein-DNA complexes (lane 12). Finally, binding of R to RRE-DR1 was also impaired by the insertions in mutants Rt356i60, Rt356i81, Rt356i184 and Rt356i233 (lanes 13 to 16). In conclusion, the DNA binding domain is contained within the first 280 N-terminal amino acids of the protein.

R binds as a dimer to DNA and forms stable dimers in solution

Since R interacts in vitro with specific DNA sequences, we next investigated whether it binds as a monomer or as a homodimer. For this, the wild-type protein and the deletion mutant Rt356 were produced by in vitro translation either alone or together. The binding activity of equimolecular amounts of the proteins was examined by electrophoretic mobility shift assay, using the RRE-DR1 probe (Figure 2A). When the two proteins were translated separately, specific retarded complexes of low (A) and high (C) mobility, corresponding respectively to the wild-type and the truncated protein-DNA complexes, were observed (lanes 3 and

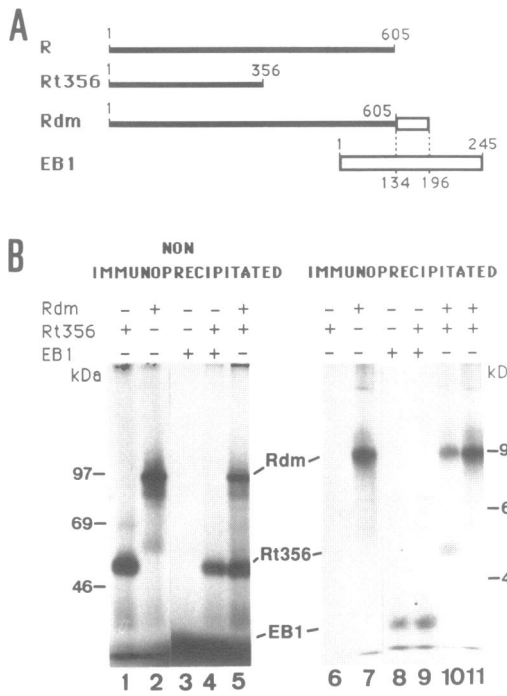

Figure 3. R forms stable dimers in solution. (A) Schematic representation of the proteins used. The Rdm protein consists of a fusion between the whole R coding frame (in black) fused with an EB1 peptide, amino acids 134 to 196 (white box). (B) SDS PAGE (8% gel) of proteins translated in vitro in the presence of ^{35}S-Methionine prior to or after immunoprecipitations. The translated proteins were either loaded onto the gel directly after in vitro translation (lanes 1 to 5) or after immunoprecipitation with the anti-EB1 antibody (lanes 6 to 11). The in vitro translated proteins are as follows: lanes 1 and 6, the truncated Rt356 mutant; lanes 2 and 7, the Rdm fusion protein; lanes 3 and 8, EB1 translated from plasmid pSPTEB1; lanes 4 and 9, Rt356 and EB1 cotranslated; lanes 5 and 10, Rt356 and Rdm cotranslated and lane 11, Rt356 and Rdm translated separately and mixed prior to immunoprecipitation.

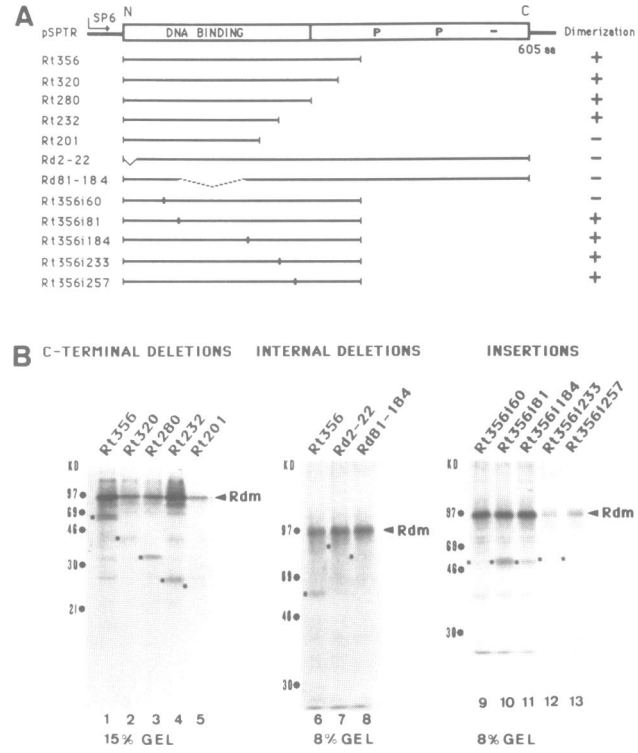

Figure 4. Localization of the dimerization domain. (A) Schematic representation of truncated and deleted R mutants that have been tested for dimerization with the Rdm fusion protein shown in Figure 3A. The mutants are named as described in the legend of Figure 1. The dimerization capacity of each mutant is indicated by + or −. (B) SDS-PAGE showing ^{35}S-Methionine labelled proteins immunoprecipitated with the anti-EB1 antibody. The R mutants were cotranslated with the Rdm fusion protein, then immunoprecipitated with the anti-EB1 antibody. Bands corresponding to Rdm are shown by an arrow on the side of the gels. Asterisks indicate the place of the size markers and the black squares locate the expected positions for the bands corresponding to each mutant, as determined on a parallel gel where the proteins were loaded directly after in vitro translation.

4, Figure 2B). However, when the two proteins were cotranslated, a third complex (B) of intermediate mobility appeared, suggesting the formation of heterodimeric structures between the full-length and the truncated proteins (lane 6). When R and Rt356 were synthesized separately and mixed before the binding reaction, no complex of intermediate mobility was found which suggests that the half life of the R and Rt356 homodimers is longer that the incubation time (lane 5).

This type of experiment however, did not demonstrate that homodimers are formed in the absence of a DNA binding site. To show that R forms stable dimers in solution, we made a construct, pSPTRdm, which allows the production of a fusion protein between R and 65 amino acids of the EB1 protein (Figure 3A), and we used a specific rabbit antibody directed against the EB1 protein moiety. This should allow the specific immunoprecipitation of the Rdm fusion protein but not of R or R mutants. As shown in Figure 3B, in vitro translated Rt356 (lane 1) was not immunoprecipitated by the anti-EB1 antibody (lane 6), whereas in vitro translated Rdm (lane 2) was immunoprecipitated (lane 7). However, when Rdm and Rt356 were cotranslated (lane 5), the two proteins were immunoprecipitated by the anti-EB1 antibody (lane 10), suggesting that Rdm and Rt356 associate in solution. This association was not observed when Rdm and Rt356 (lanes 1 and 2) were synthesized separately and mixed before incubation with the anti-EB1 antibody since, in that case, only the Rdm protein was immunoprecipitated (lane 11). The immunoprecipitation of Rt356 with Rdm was also not due to an interaction between the EB1 protein and R. In effect, when Rt356 was cotranslated with EB1 (lane 4), only EB1 was immunoprecipitated by the anti-EB1 antibody (lane 9) as compared to the EB1 polypeptide precipitated alone (lane 8). No immunoprecipitation of proteins was observed using a non immune antiserum (not shown).

In conclusion, both the results of the electrophoretic mobility shift assays and of the immunoprecipitation experiments, suggest that R binds to DNA as a homodimer and that dimerization can occur in absence of DNA.

Localization of the dimerization region

The experiments described above suggest that R can form homodimers in solution and that the first 356 N-terminal amino acids contain the putative dimerization region. This region overlaps with the domain necessary for binding to DNA. To further localize the region of the protein necessary for dimerization, we cotranslated Rdm and the R mutant proteins presented in Figure 4A and made use of the selective immunoprecipitation assay described above.

We first tested the capacity of C-terminal truncated proteins to dimerize with Rdm. The C-terminal truncated proteins Rt356,

Rt320, Rt280 and Rt232 were coimmunoprecipitated with the Rdm fusion protein (Figure 4B, lanes 1 to 4). However, this was not the case for the shortest protein Rt201 (lane 5) indicating that the region between amino acids 232 and 201 is necessary for dimerization.

At the N-terminal end of the protein, deletion of the first 20 amino acids (mutant Rd2−22) also impaired dimerization (lane 7) demonstrating that the dimerization domain is located within the first 232 amino acids of R. Other mutants within the N-terminal part of the protein were also tested. The deletion in the centre of the 232 N-terminal amino acids (mutant Rd81−184) affected dimerization (Figure 3B, lane 8).

One insertion in the N-terminal part of the protein (mutant Rt356i60) also impaired dimerization of the protein (Figure 3B, lane 9). However, more interestingly, insertions in the centre of the 232 N-terminal amino acids of the protein (mutants Rt356i81 and Rt356i184) did not affect dimerization (Figure 4B, lanes 10 and 11), although these insertions completely impaired DNA binding (Figure 1C, lanes 14 and 15). Accordingly, insertions outside of the minimal 232 amino acids necessary for dimerization (mutants Rt356i233 and Rt356i257) did not affect dimerization (Figure 4B, lanes 12 and 13).

In conclusion, the analysis of both the DNA binding properties and the dimerization capacities of a series of deletion and insertion mutants of the R protein showed that the regions necessary for these two functions are overlapping. Furthermore, each mutant which failed to dimerize did not bind to DNA. This suggests that dimerization is a prerequisite for DNA binding. It seems nevertheless, that discrete regions within the domain are specifically responsible for the DNA binding function.

The C-terminal part of the R protein contains regions required for transcriptional activation

As shown above, R activates specific transcription probably by direct interaction with specific DNA sequences. Since the first 280 N-terminal amino acids are necessary and sufficient for homodimerisation and binding to specific DNA sequences, the region necessary for transcriptional activation is likely to be localized in the C-terminal part of the protein. To test this possibility, we generated a number of in frame deletion and insertion mutants in this region (Figure 5A) and assayed their activating potential. All the proteins used in this study, were first compared for stability and nuclear localization. This was done by SDS-PAGE and immunoblotting, after expression of R and R mutant proteins in COS-7 cells and separation of the nuclear fraction from the cytoplasmic fraction. Immunoblots were incubated with a rabbit antibody raised against the 572 amino acid C-terminal of the R protein (35) (not shown). Only mutant proteins whose stability and repartition between nuclear and cytoplasmic fractions were similar to that observed for the wild-type R protein, were used.

The wild-type R protein, expressed under the control of the SV40 early promoter from plasmid pKSVR (29), has previously been shown to activate specific transcription from plasmid pG1B0 (1). This construct contains the RRE-DR1 placed upstream of the β-globin promoter (Figure 5B). In order to evaluate the transcriptional activation potential of the R mutants, pG1B0 was transfected into HeLa cells either alone or together with plasmids expressing R or the various R mutants described in Figure 5A. The level of specific transcription from plasmid pG1B0 was evaluated by S1 nuclease digestion.

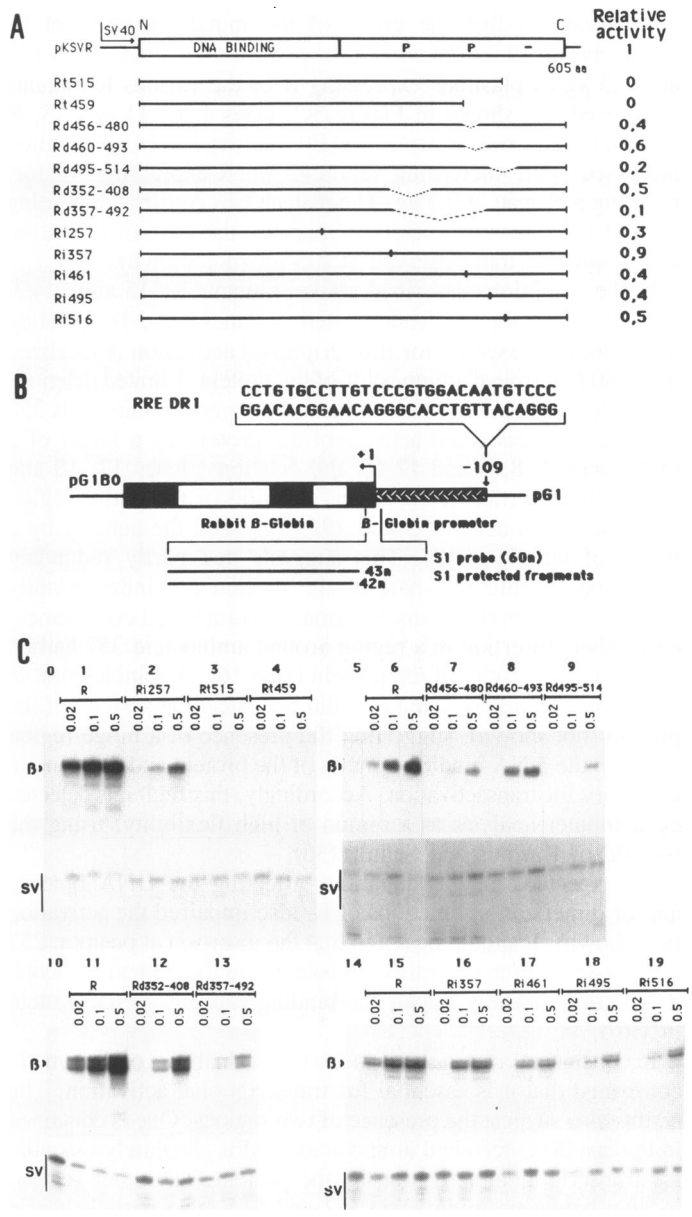

Figure 5. Mapping of regions necessary for transactivation. (A) Schematic representation of truncated and deleted R mutants used in the transactivation assays. The mutants are named following the rules edicted in Figure 1. All, including the wild-type R were cloned under the dependance of the SV40 early promoter, in the pKSV10 expression vector. The relative ability of each mutant to transactivate is expressed as a proportion of the wild type activity. The respective activities are evaluated as described below. (B) Schematic representation of plasmid pG1B0. The double stranded oligonucleotide **TGTGCC**TTGTCCC**GTGGAC**AATGTCC containing the RRE-DR1 represented in bold letters, was cloned 109 bp upstream from the transcription start site of the β-globin reporter gene. The S1 probe and the fragments resulting of S1 nuclease digestion are indicated. (C) Hela cells were transfected with 5 μg of β-globin reporter gene plasmid together with increasing amounts (20, 100 and 500 ng) of R- or various R mutants-expressing vector. Transcriptional activation was determined by quantitative S1 analysis of total cellular RNA isolated from transfected cells. The specific β-globin bands are indicated by β in the margin. Plasmid pSV2β expressing an SV40-β-globin hybrid RNA under the control of the SV40 early promoter-enhancer was cotransfected as an internal control. The specific bands are indicated by SV in the margin. The level of activation was quantified by cutting out the bands and counting the amount of radioactivity in each band. The results are expressed as a fraction of the activation level for R in the same experiment. The numbers shown in part A of the figure are those calculated for 0.1 μg of transfected R- or R mutants-expressing vector.

To ascertain that the effect of the mutations will not be counterbalanced by the amount of protein expressed, 0.02, 0.1 and 0.5 μg of plasmids expressing R or the various R mutants were used. As shown in Figure 5C, lanes 1, 6, 11 and 15, R activates transcription from pG1B0 and the level of activation increased with increasing amounts of R-expressing vector, reaching a plateau at 0.5 μg. The plateau was confirmed by using 2 μg of R-expressing vector. In this case the level of activation is the same as that observed at 0.5 μg (not shown).

In the conditions described above, mutants Rt515 and Rt459 completely lost their activation potential (lanes 3 and 4) indicating that a domain essential for transcriptional activation is localized in the 90 C-terminal amino acids of the protein. Limited deletions or single insertions in a region located between amino acids 352 and 515, decreased the activity of the protein by a factor of 2 to 4 (lanes 7, 8, 9 and 12 for the deletions; lanes 17, 18 and 19 for the insertions). However, deletion of the entire region between amino acids 352 and 492 decreased the activity by a factor of ten (lane 13). This suggests that partly redundant subregions within this part of the protein are independantly capable of conferring transcriptional activation and complement each other. Insertion in a region around amino acid 357 had no effect on the activity of the protein (lane 16). A minideletion in this region (5 amino acids) also did not affect the activity of the protein (not shown), suggesting the presence of a hinge region between the DNA binding domain of the protein and the domain necessary for transactivation. Accordingly, this region is detected by computer analysis as a region of high flexibility, using the method of Karplus and Schulz (36).

As expected, all the mutations affecting the DNA binding and/or dimerization functions of R, also impaired the activation (not shown). It should be noted that the insertion at position 257 led to a decrease in the activation potential of the protein, although it did not noticeably impair the binding capacity of the protein in vitro.

In conclusion, mutagenesis in the C-terminal part of the protein confirmed that it is essential for transcriptional activation. The results also suggest the presence of two regions. One is contained in the last 90 C-terminal amino acids and is absolutely essential for transcriptional activation by the protein. The other region, located between amino acids 352 and 515, also participates in the transcriptional activity but can be deleted without a complete loss of the function.

DISCUSSION

We have established that the active form of the EBV transcription factor R, appears to be a homodimer as determined by its capacity to form specific complexes with a DNA sequence in vitro. Such homodimers can also form in solution in the absence of DNA and are relatively stable. Furthermore, although R/Rt356 heterodimers were formed and bound to DNA in vitro when the two proteins were cotranslated, this did not occur when the two proteins were mixed just before incubation with the DNA. In this case only DNA-bound complexes corresponding to R/R homodimers and Rt356/Rt356 homodimers were observed, suggesting that the dissociation of the homodimers is too slow to allow protein exchange during the time of the experiments (30 minutes).

The analysis of the binding properties of R mutants shows that 320 amino acids at the N-terminus of the protein are sufficient for specific DNA recognition. This region also contains the structure(s) necessary for dimerization. In the case of R, dimerization appears to be a prerequisite for binding, at least in vitro: all the mutants that failed to dimerize also failed to bind to DNA. The manner by which the R homodimers interact with the DNA is unknown. We have previously shown that R simultaneously contacts two core sequences separated by 5 to 7 bp (1). The binding of R as a homodimer and the fact that the binding sites are only partially palindromic raises different possibilities to explain how the R homodimers interact with DNA. Either each component of the homodimer contacts each of the two core sequences or just one component of the dimer contacts the DNA. Whether the formation of dimers is necessary in vivo for DNA binding remains to be determined. There is also the possibility that R could associate with cellular proteins to form heterodimers in vivo.

The region necessary for dimerization within the DNA binding region, can be limited to the first 232 N-terminal amino acids. This implies that amino acids between 232 and 320 are only concerned with DNA binding and are not required for dimerization. The same can be said for amino acid sequences at the centre of the DNA binding/dimerization region,.as some insertion mutants completely impaired DNA binding but did not affect dimerization. From these observations, it seems that the dimerization region and the DNA binding domain per se are probably more overlapping than juxtaposed as is the case for other DNA binding proteins such as the leucine zipper family of transcription factors (For a review see ref. 37).

It is difficult, with such a large region implicated in both dimerization and DNA binding, to assign specific structures within this region, that are directly involved in protein:protein or protein:DNA interactions. Moreover, no obvious similarities to the previously characterised motifs of various DNA binding proteins (38) could be detected. We could not therefore, on the basis of sequence homology alone, identify the precise segment(s) of R responsible for dimerization or DNA binding. It is thus possible that R specifically recognises its binding site by using a novel type of DNA binding/dimerization structure.

From the analysis of the transactivation efficiency of the different R mutants, two regions, one rich in prolines (amino acids 352 to 515), the other rich in acidic residues (amino acids 515 to 605), seem to be important for this function. Within the first region, the prolines are organised into two clusters. The first cluster, localized between amino acids 352 and 410 contains 22% proline, the second cluster which is between amino acids 450 and 500 contains 34% proline. Deletions of either cluster 1 (mutant Rd352-408) or cluster 2 (mutants Rd456-480 or Rd460-493), only decreased the activation potential of R by approximately a factor 2, while deletion of both clusters (mutant Rd357-392) reduced transactivation by a factor of ten. These results suggest that the two proline-rich clusters can complement each other. Such proline-rich regions involved in activation has been found in other transcription factors, in particular CTF/NF1 (39).

The second region is very acidic which is a characteristic shared with the activating regions of other transcriptional activating factors (40; 41). This domain seems essential to the activation property of the R protein: when the acidic region in the C-terminal part of the protein is deleted, the proline-rich region has no activity, suggesting that both domains act cooperatively. Whether this is the case will be determined by evaluating the transcriptional potential of the two regions linked individually or in combination to an heterologous DNA binding domain.

Transactivation domains are thought to interact with components of the basic transcription complex, either directly or via auxilliary factors (see ref. 42 for a review).In the case of R, one of these putative cellular factors seems to also interact, directly or indirectly, with the other transcription factor EB1. In this regard, we have shown elsewhere (43) that EB1 is able to repress the R-activated transcription, either in the absence of an EB1 binding site on the reporter promoter or in the absence of a functional EB1 DNA binding domain. By these properties this inhibitory effect is related to 'squelching' (44). Unbound EB1 would thus sequester a cellular factor required for R to contact the basal transcription apparatus, resulting in repression of R-induced transcription. Domain(s) of the R protein involved in the interaction with such a cellular factor are not yet identified. Understanding of the interactions between EB1, R and this cellular factor will depend upon purified fractions and in vitro assays.

In conclusion, we have been able to ascribe several functions to separate regions of the R protein. However, the biochemical characterization of the dimerization and DNA binding structures has yet to be established.

ACKNOWLEDGEMENTS

We wish to thank H. Wolf and G. Miller for the gift of the R and Z antisera, P. Jalinot for plasmid pG1 and Conrad B. Bluink for reading the manuscript. This work was financially supported by the Federation Nationale des Centres de Lutte contre le Cancer (F.N.C.L.C.C.), by I.N.S.E.R.M. (contrat no 871015), and by the Association pour la Recherche sur le Cancer (A.R.C., no 6810). H.G. and J.F.G. are recipients of MRT fellowships.

REFERENCES

1. Gruffat, H., Manet, E., Rigolet, A. and Sergeant, A. (1990) Nucleic Acids Res., 18, 6835–6843.
2. Lerner, M. , Andrews, N., Miller, G. and Steitz, J. (1981) Proc. Natl. Acad. Sci. USA, 78, 805–809.
3. Miller, G. (1985) In B. Fields (ed.), Virology, Raven Press, Publishers, New York, 563–590.
4. Bodescot, M., and Perricaudet, M. (1986) Nucleic Acids Res., 14, 7103–7114.
5. Pfitzner, A. J., Strominger, J. L. and Speck, S. H. (1978) J. Virol., 61, 2943–2946.
6. Hennessy, K. , Fennewald, S., Hummel, M., Cole, T. and Kieff., E. (1984) Proc. Natl. Acad. Sci. USA, 81, 7207–7211.
7. Laux, G. , Perricaudet, M. and Farrell, P. J. (1988) EMBO J., 7, 769–744.
8. Sample, J., Liebowitz, D. and Kieff, E. (1989) J. Virol., 63, 933–937.
9. Zur Hausen, H. , J. O'Neil, E. , Freese, U. K. and Hecher, E. (1978) Nature (London), 272, 373–375.
10. Miller, G., Rabson, M. and Heston, L. (1984) J. Virol., 50, 174–182.
11. Countryman, J. and Miller, G. (1985) Proc. Natl. Acad. Sci. U. S. A., 82, 4085–4089.
12. Countryman, J. K., Jenson, H., Grogan, E. and Miller, G. (1986) Cancer Cells, 4, 517–523.
13. Chevallier-Greco, A. , Manet, E., Chavrier, P., Mosnier, C., Daillie, J. and Sergeant, A. (1986) EMBO J., 5, 3243–3249.
14. Takada, K., Shimizu, N., Sakuma, S. and Ono, Y. (1986) J. Virol., 57, 1016–1022.
15. Chevallier-Greco, A., Manet, E., Chavrier, P., Urier, G., Buisson, M. Daillie, J. and Sergeant, A. (1987) In D. L. Ablashi (ed), Epstein-Barr virus and Human diseases. The Humana Press, Clifton, N. J., 157–161.
16. Hardwick, J. M., Lieberman, P. M. and Hayward, S. D. (1988) J. Virol., 62, 2274–2284.
17. Farrell, P. J., Rowe, D. T., Rooney, M. C. and Kouzarides, T. (1989) EMBO J., 8, 127–132.
18. Lieberman, P. M. and Berk, A. J. (1990) J. Virol., 64, 2560–2568.
19. Chang, Y. N., Dong, D. L. Y., Hayward, G. S. and Hayward, S. D. (1990) J. Virol., 64, 3358–3369.
20. Urier, G. , Buisson, M., Chambard, P. and Sergeant, A. (1989) EMBO J., 8, 1447–1453.
21. Flemington, E. and Speck, S. H. (1990) J. Virol., 64, 1227–1232.
22. Chevallier-Greco, A., Gruffat, H., Manet, E., Calender, A. and Sergeant, A. (1989) J. Virol., 63, 615–623.
23a. Kenney, S., Holley-Guthrie, E., Mar, E. C. and Smith M. (1989) J. Virol., 63, 3878–3883.
23b. Kenney, S., Kamine, J., Holley-Guthrie, E., Mar, E. C., Lin, J. C., Markovitz, D. and Pagano, J. (1989). J. Virol., 63: 3870–3877.
24. Buisson, M., Manet, E., Biemont, M. C., Gruffat, H., Durand, B. and Sergeant, A. (1989) J. Virol., 63, 5276–5284.
25. Cox, M. Leahy, J. and Hardwick, M. (1990) J. Virol., 64, 313–321.
26. Gruffat, H., Moreno, N. and Sergeant, A. (1990) J. Virol., 64, 2810–2818.
27. Laux, G., Freese, U. K. and Bornkamm, G. (1985) J. Virol., 56, 987–995.
28. Hammerschmidt, W., and Sugden, B. (1988) Cell, 55, 427–433.
29. Manet, E., Gruffat, H., Trescol-Biemont, M-C., Moreno, N., Chambard, P., Giot, J. F. and Sergeant, A. (1989) EMBO J., 8, 1819–1826.
30. R. D. Everett. (1987) EMBO J., 6, 2069–2076.
31. Norman, C., Runswick, M., Pollock, R. and Treisman, R. (1988) Cell, 55, 989–1003.
32. Taylor, N., Countryman, J., Rooney, C., Katz, D. and Miller, G. (1989) J. Virol., 63, 1721–1728.
33. Graham, F. L. and Van Der Eb, A. J. (1973) Virology, 52, 456–467.
34. Jalinot, P. and Kedinger, C. (1986) Nucleic Acids Res., 14, 2651–2669.
35. Seibl, R., Motz, M. and Wolf, H. (1986) J. Virol., 60, 902–909.
36. Karplus, P. A. and Schulz, G. E. (1985) Naturwissenschaften, 72, 212–213.
37. Busch, S. J. and Sassone-Corsi, P. (1990) TIG, 6, 36–40.
38. Johnson, P. F. and McKnight, S. L. (1989) Annu. Rev. Biochem., 58, 799–839.
39. Mermod, N., O'Neill, E. A., Kelly, T. J. and Tijan, R. (1989), Cell, 58, 741–753.
40. Hope, I. A., Mahadevan, S. and Struhl, K. (1988) Nature, 333, 635–640.
41. Ma, J. and Ptashne, M. (1987) Cell, 48, 847–853.
42. B. Lewin (1990), Cell, 61, 1161–1164.
43. Giot, J. F., Mikaelian, I., Buisson, M., Manet, E., Joab, I., Nicolas, J. C. and Sergeant, A. (1991) Nucleic Acids Res., In press.
44. M. Ptashne (1988) Nature, 335, 683–689.