10-1-1998

Mutational analysis of a transcriptional activation region of the VP16 protein of herpes simplex virus

Susan M. Sullivan
Michigan State University

Peter J. Horn
Michigan State University

Victoria A. Olson
Michigan State University

Allen H. Koop
Calvin University

Nucleic Acids Research, 1998, Vol. 26, No. 19 **4487–4496**

Mutational analysis of a transcriptional activation region of the VP16 protein of herpes simplex virus

Susan M. Sullivan, Peter J. Horn, Victoria A. Olson+, Allen H. Koop[1], Wei Niu[2], Richard H. Ebright[2] and Steven J. Triezenberg*

Department of Biochemistry, Michigan State University, East Lansing, MI 48824-1319, USA, [1]Department of Biology, Calvin College, Grand Rapids, MI 49546, USA and [2]Howard Hughes Medical Institute, Waksman Institute and Department of Chemistry, Rutgers University, New Brunswick, NJ 08855, USA

Received May 11, 1998; Revised August 5, 1998; Accepted August 12, 1998

ABSTRACT

The VP16 protein of herpes simplex virus is a potent transcriptional activator of the viral immediate early genes. The transcriptional activation region of VP16 can be divided into two functional subregions, here designated VP16N (comprising amino acids 413–456) and VP16C (amino acids 450–490). Assays of VP16C mutants resulting from both random and alanine-scanning mutagenesis indicated that the sidechains of three phenylalanines (at positions 473, 475 and 479) and one acidic residue (glutamate 476) are important for transcriptional activation. Aromatic and bulky hydrophobic amino acids were effective substitutes for each of the three Phe residues, whereas replacement with smaller or polar amino acids resulted in loss of transcriptional function. In contrast, many changes were tolerated for Glu476, including bulky hydrophobic and basic amino acids, indicating that the negative charge at this position contributes little to the function of this subregion. Similar relative activities for most of the mutants were observed in yeast and in mammalian cells, indicating that the structural requirements for this activation region are comparable in these two species. These results reinforce the hypothesis that bulky hydrophobic residues, not acidic residues, are most critical for the activity of this 'acidic' transcriptional activation region.

INTRODUCTION

Basal levels of transcription of eukaryotic mRNA genes by RNA polymerase II can be augmented by the action of proteins known as transcriptional activators. Several models have been proposed for the function of eukaryotic activators, including the recruitment of chromatin remodeling enzymes (1) or of basal transcription factors. These latter interactions might facilitate pre-initiation complex formation (2), increase initiation rate or aid in promoter clearance (3) or increase the rate of elongation (4). Each of these models invokes specific interactions between transcriptional activation regions and particular target proteins that are likely to require certain structural motifs present in the activator protein. Therefore, transcriptional activators have been studied intensively to identify such activation motifs.

Eukaryotic transcriptional activators were initially described by the amino acids most abundantly present, resulting in classes of acidic, glutamine-rich, proline-rich and serine/threonine-rich regions (3,5), although some activation regions do not fit this scheme (6). Thorough mutational analyses of acidic activation regions from VP16 (7,8), RelA (9), p53 (10), Bel1 (11), Gcn4 (12), the glucocorticoid receptor (13,14), C1 (15) and of a glutamine rich activation region from Sp1 (16), indicated that the most prevalent types of amino acid were not those most critical for activation function. Rather, in each case, specific bulky hydrophobic or aromatic amino acids were more important.

This conclusion from mutational analyses is consistent with results from biophysical experiments. Although activation regions in isolation typically appear highly unstructured (17–21), more ordered structures can be induced by the presence of putative target proteins. Among the examples reported to date are the interaction of the VP16 activation region with TBP (22) and with $hTAF_{II}32$ (23), the interaction of the p53 transcriptional activation region with its repressor MDM2 (24), and the interaction of an activation region from CREB with a fragment of CBP (25). In each case, the interaction surface of the activator comprised primarily hydrophobic residues (including many defined by mutational analysis) within helical segments, with a notable lack of participation by acidic sidechains.

VP16 (also known as Vmw65 or α-TIF) is a component of the virion of herpes simplex virus whose transcriptional role is to activate expression of the viral immediate early genes (reviewed in 26). The 490 amino acids of VP16 can be divided into two domains by biochemical or molecular genetic methods (27,28). A region comprising the N-terminal 410 amino acids forms a complex with the host proteins Oct-1 and HCF that can bind to specific DNA sequences in the promoters of the immediate early genes (26). The C-terminal 80 amino acids of VP16 are both necessary and sufficient for transcriptional activation, even when fused to a heterologous DNA-binding region (27,29,30). The

*To whom correspondence should be addressed. Tel: +1 517 353 7120; Fax: +1 517 353 9334; Email: triezenb@pilot.msu.edu

+Present address: Department of Biochemistry, University of Wisconsin, Madison, WI 53706, USA

VP16 activation region can be further divided into two subregions herein termed VP16N (amino acids 412–450) and VP16C (amino acids 450–490), each of which is independently capable of activation when provided with a DNA binding mechanism (8,27,31–33).

Extensive mutational analysis of the VP16N subregion indicated a key role for the aromatic character of F442 and supporting roles for other bulky hydrophobic residues, specifically L439 and L444 (7,8). The acidic residues in this region had a much less important role. Any one or even several negative charges could be removed without significant loss of function. We and others have previously tested the roles of several amino acids within the C-terminal subregion of VP16. In one case, several substitutions of F473 and F475 were tested in the context of full-length VP16 bearing the debilitating F442A mutation in the N subregion, with the result that mutations at F475 had a somewhat greater effect than corresponding mutations at F473 (8). In another study, the simultaneous mutation of three phenylalanine residues within the C subregion (i.e. F473, F475 and F479) to alanine had a dramatic effect on the ability of that subregion to function in the context of the full-length activation region (32).

These preliminary analyses, however, are insufficient to ensure that all key amino acids in VP16C have been identified, and fall short of defining the chemical characteristics of those amino acid sidechains that contribute to the transcriptional activity of this region. To comprehensively identify the critical residues in VP16C and to further test the hypothesis that specific patterns of hydrophobic residues are important for function of activation regions, we have undertaken a thorough mutational analysis of VP16C using both PCR-mediated random mutagenesis and systematic alanine-scanning mutagenesis. Random mutagenesis permits the analysis of a broad range of amino acid substitutions altering the size, charge and hydrophobicity of the amino acid sidechains. Moreover, the approach as applied here has the advantage of an *in vivo* screening step which allows selection of mutants defective in transcriptional activation. The complementary strategy of alanine-scanning mutagenesis ensures that each amino acid position will be tested and indicates the contributions of the sidechain atoms of a given residue, because alanine substitutions eliminate all sidechain atoms beyond the β-carbon (34). The combination of these complementary mutational analyses, therefore, provides information about each position while focusing on those that affect activation most strongly. Our results indicate that in the case of VP16C, as with other activation regions, specific hydrophobic amino acids are particularly critical to its transcriptional function.

MATERIALS AND METHODS

Plasmids and phage

The substrate for all mutagenesis reactions was mpPJH18, an M13mp19 phagemid containing a *Sma*I–*Bam*HI fragment encoding VP16 residues 452–490 with 119 bp of 3′ untranslated sequence. Mutants generated by error-prone PCR were screened in yeast using pPJH13, a high copy (2μ *LEU2*) plasmid that expresses a Gal4(1-147)–VP16C fusion protein from the *ADH1* promoter. Quantitative assays of transcriptional activation in yeast were performed using wild type or mutant versions of pVS1, a low copy plasmid (*CEN6 ARSH4 LEU2*) that expresses the Gal4(1-147)–VP16C fusion protein, and the yeast reporter construct pLGSD5 (35) containing the *Escherichia coli lacZ* structural gene with the *GAL1-10* upstream activating sequence near the *CYC1* promoter. VP16C mutants were also assayed in mammalian cells using pSM71-1.VP16C, a derivative of pSGVP (29) that expresses the Gal4(1-147)–VP16C fusion protein from the SV40 early-gene promoter. The mammalian reporter plasmid pG_5BCAT (36) expresses the bacterial chloramphenicol acetyltransferase protein under the control of five Gal4 sites upstream of the TATA sequence from the adenovirus E1B promoter.

Error-prone PCR-mediated mutagenesis

To generate random nucleotide substitutions, the VP16C coding region of mpPJH18 was amplified by PCR using forward and reverse M13 sequencing primers under reaction conditions that diminish the accuracy of *Taq* DNA polymerase, including the presence of $MnCl_2$ (37) and limiting concentrations of dATP (38). Reaction mixtures included 10 mM Tris–Cl (pH 8.8), 50 mM KCl, 0.5 mM $MgCl_2$, 0.5 mM $MnCl_2$, 0.4 mM dATP, 3.2 mM each dCTP, dGTP, dTTP, 100 μg/ml gelatin, 10 ng of mpPJH18, 60 pmol of each primer, and 5 U of *Taq* DNA polymerase (Perkin-Elmer) in a final volume of 100 μl. After an initial melting step of 2 min at 94°C, the temperature was lowered to 80°C and PCR was initiated by the addition of nucleotides and polymerase. Amplification proceeded through 25 cycles (30 s annealing at 43°C, 2 min elongation at 72°C, 30 s denaturation at 94°C). The PCR products were digested with *Sal*I and *Bam*HI, ligated into pPJH13 and transformed into yeast strain BP1 (*MAT*a *ura3-52 leu2-3, 2-112 Gal4::HIS4 ade1-100*; 39).

Alanine scanning and site-directed mutagenesis

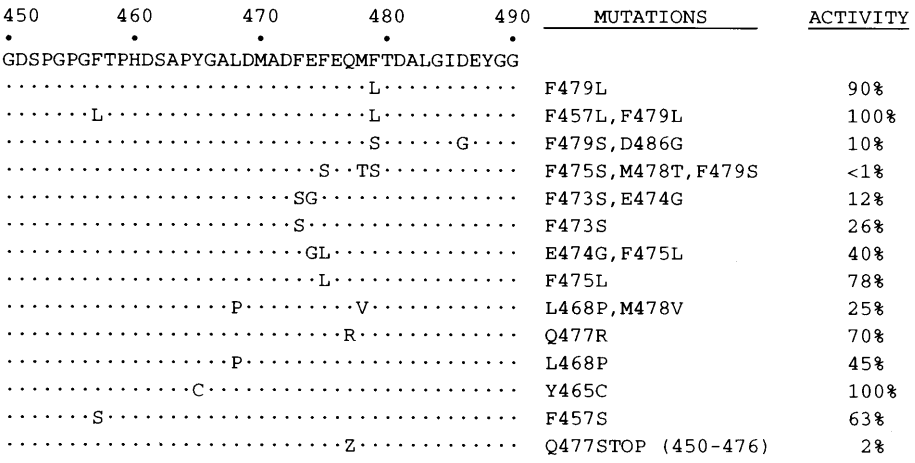
Oligonucleotide-directed mutations were generated in mpPJH18 using the method described by Kunkel (40). In some cases, a variety of substitutions at specific positions were generated using degenerate oligonucleotides. *Sal*I–*Bgl*II fragments containing the mutations were subcloned into pVS1 and resequenced to verify the identities of mutations.

Yeast transformation and β-galactosidase assay

Plasmids expressing VP16C mutants fused to the Gal4 DNA-binding region were co-transformed with the reporter plasmid pLGSD5 into yeast strains BP1 or PSY316 (*MAT*a *ade2-101 his3-del.200 leu2-3,2-112 lys2 ura3-53*; 39) using the procedure of Gietz *et al.* (41). To assay the activity of β-galactosidase produced from pLGSD5, 2 ml cultures containing pools of ~100 yeast transformant colonies were grown to early stationary phase (OD_{600} = 0.9–1.2) in selective media, then diluted to OD_{600} = 0.1, grown to mid-log phase (OD_{600} = 0.5) and harvested by centrifugation. β-galactosidase activity was measured (42) using extracts obtained from 5 ml cultures with results normalized to protein concentration as determined by Bradford assay (43).

Yeast immunoblot analysis

To assess the steady-state levels of each mutant activator protein, 25 ml yeast cultures grown as described above were harvested by centrifugation, washed once in water and resuspended in water to a total volume of 100 μl. An equal volume of 4× SDS–PAGE loading buffer lacking bromphenol blue was added and extracts were boiled for 15 min, then centrifuged for 5 min at 14 000 *g* to clarify the extract. Relative protein concentrations for each extract were determined using a detergent compatible protein



450 460 470 480 490	MUTATIONS	ACTIVITY
GDSPGPGFTPHDSAPYGALDMADFEFEQMFTDALGIDEYGG		
·····L·····	F479L	90%
·····L·····L·····	F457L,F479L	100%
·····S·····G····	F479S,D486G	10%
·····S··TS·····	F475S,M478T,F479S	<1%
·····SG·····	F473S,E474G	12%
·····S·····	F473S	26%
·····GL·····	E474G,F475L	40%
·····L·····	F475L	78%
·····P·····V·····	L468P,M478V	25%
·····R·····	Q477R	70%
·····P·····	L468P	45%
·····C·····	Y465C	100%
·····S·····	F457S	63%
·····Z·····	Q477STOP (450-476)	2%

Figure 1. Mutation identities and relative activities of VP16C mutants selected for reduced toxicity when highly expressed in yeast. The Gal4–VP16C fusion protein, when expressed from the *ADH1* promoter on a high-copy plasmid, is toxic to yeast resulting in tiny colonies. The VP16C mutants listed here, generated by error-prone PCR, were selected based on larger yeast colony size. The VP16C subregions were then recloned into a low-copy vector (to avoid toxicity) and assayed for transcriptional activity using a Gal4-reponsive *lacZ* reporter plasmid (pLGSD5). Below the amino acid sequence of VP16C (residues 450–490) are shown the positions and types of substitutions identified in each of these mutants. The β-galactosidase activities in lysates of cells containing each of the Gal4-VP16C mutants are expressed relative to the activity of the wild type Gal4–VP16C fusion protein.

assay (Bio-Rad), and ~10 µg total protein were loaded per sample onto 15% SDS–PAGE gels. Following electrophoresis, proteins were electrophoretically transferred to nitrocellulose. Membranes were blocked in 10% powdered milk, 20 mM Tris–Cl pH 7.5, 137 mM NaCl, 4 mM KCl, 0.01% Tween 20 and then incubated sequentially with rabbit polyclonal antibodies raised against the Gal4–VP16 fusion protein (LA2-3; L.Alexander and S.J.Triezenberg, unpublished) followed by goat anti-rabbit IgG conjugated to horseradish peroxidase (Bio-Rad). Blots were developed using an enhanced chemiluminescence system (ECL, Amersham or Renaissance, DuPont NEN Life Science Products).

Mammalian plasmid construction and assay

The VP16C mutant activation regions were subcloned from the yeast plasmids into pSM71-1.VP16C. All subclones were sequenced to verify their identities. 100 ng of each plasmid were co-transfected into mouse L cells (American Type Culture Collection) with 1 µg pG_5BCAT using DEAE–dextran (44). Plates were incubated at 37°C, 10% CO_2 for 36–48 h, after which cells were harvested by scraping into phosphate-buffered saline followed by centrifugation. Cell extracts were made by freeze–thawing and CAT assays were performed by the mixed-phase method (45,46). Final reaction volumes of 100 µl contained 30 µl of cell extract, 200 mM Tris–HCl pH 7.8, 4 mM EDTA, 30 mM acetyl-CoA and 0.4 µCi [^{3}H]acetyl CoA. Reactions were overlaid with 10 ml Econofluor2 (Packard Industries), incubated for 2 h, and quantitated using an LKB Wallace 1209 Rackbeta liquid scintillation counter. Counts were normalized to protein concentration as determined by Bradford assay (43).

Mammalian western blot analysis

Separate plates of L cells were transformed with 5 µg of each plasmid encoding wild type or mutant Gal4–VP16 fusion protein in the manner described above. After 36–48 h, cells were washed once in phosphate-buffered saline and harvested by scraping into 200 µl of 4× SDS–PAGE loading buffer lacking bromphenol blue. Samples were boiled for 10 min and then spun for 5 min at 14 000 *g* to clarify the extracts. Relative protein concentrations were determined by using the Bio-Rad detergent-compatible protein assay and western blots were performed as described above.

RESULTS

Biological selection of transcriptionally defective mutants of VP16C

Expression of the wild type Gal4–VP16C fusion protein from a high-copy plasmid is relatively toxic to yeast, as is the fusion protein bearing the full-length VP16 activation region (39). As a result, yeast strains expressing high levels of these activators produce very small colonies. Mutations in the VP16 activation region that reduce transcriptional activation also reduce the toxicity (39), yielding larger yeast colonies. This phenomenon was employed to screen a library of VP16C clones to identify mutations that negatively affect transcriptional activity. A DNA fragment encoding the VP16C subregion was amplified using *Taq* DNA polymerase under error-prone conditions. The population of amplified fragments was then ligated into a yeast expression vector such that the VP16 coding sequences were in frame with sequences encoding the Gal4 DNA-binding domain. This library was then transformed into yeast cells and large colonies were selected. DNA fragments encoding the VP16C subregion were isolated from these colonies and were sequenced. These fragments were then cloned into a low-copy yeast plasmid from which expression of the Gal4–VP16C fusion protein is sufficiently low to avoid the toxicity. Transcriptional activation by the mutant fusion proteins was quantitatively assessed using a β-galactosidase reporter gene with a Gal4-responsive promoter (Fig. 1).

Of the 14 PCR-generated mutants in VP16C isolated by this approach, 11 have at least one amino acid substitution between positions 470 and 480. Eight of the 13 missense mutants have changes at one or more of the phenylalanines in this region (F473, F475 and F479). Moreover, all mutants displaying ≤40% of wild type activity bear a substitution at one or more of these

phenylalanine residues, with the exception of the L468P/M478V mutant. For two of these three phenylalanines, substitution to leucine resulted in only a relatively modest defect: in β-galactosidase assays, F475L and F479L displayed 78 and 90% of wild type activity, respectively. Substitutions at these positions to a hydrophilic residue such as serine, however, resulted in large defects. The mutant F473S retained only 26% of wild type activity, and the two mutants with serine substitutions at F479 (F479S/D486G and F475S/M478T/F479S), showed ≤10% of wild type levels. These results indicate the important role of bulky hydrophobic or aromatic groups at positions 473, 475 and 479 for transcriptional activation by VP16C.

Interestingly, the fourth phenylalanine residue in VP16C, F457, seems much less important. This position was altered in only two of the mutants isolated by this approach. The F457L/F479L mutant activated the reporter gene at nearly wild type levels. The F457S mutant showed 63% of wild type activity, a modest reduction compared to serine substitutions at the other phenylalanine positions. This result supports the idea that residues between 470 and 480, not simply any bulky hydrophobic or aromatic residues in the vicinity, were primarily responsible for the transcriptional activity of VP16C.

Alanine scanning mutagenesis of VP16C

As a complement to the random mutagenesis and to more systematically assess the importance of each residue of VP16C individually, we constructed alanine substitutions at each position (except those that were originally alanine or glycine). The 27 mutants created in this set were tested as Gal4–VP16C fusion proteins, expressed from a low-copy plasmid, for the ability to activate a Gal4-responsive β-galactosidase reporter gene in yeast. The results (Fig. 2) support the conclusion derived from the PCR-generated mutants that the most critical residues for activation byVP16C lie between positions 470 and 480. Only four of the 27 alanine mutants have activity <70% of that of wild type. Three of these mutations affect phenylalanines: F473A (51% activity), F475A (21% activity) and F479A (43% activity). A double mutant with alanine substitutions at both positions 473 and 475 showed ≤1% activity. The fourth mutant with a significant loss of activity was E476A, which displayed only 38% of wild type activity. This result suggests that the sidechain of E476 contributes to the function of VP16C, despite the fact that this site was not affected in any of the randomly-generated mutants isolated in the previous experiment. All other alanine substitutions had little or no effect on transcriptional activation in yeast assays.

To determine whether activation by VP16C in mammalian cells depends on the same amino acids as in yeast, the alanine substitution mutants were also assayed as Gal4 fusion proteins in mouse L cells. Plasmids expressing the Gal4–VP16C proteins from the SV40 early promoter were cotransfected with pG5Bcat, a reporter plasmid in which expression of chloramphenicol acetyltransferase (CAT) is controlled by Gal4-binding sites (36). After 36–48 h, the CAT enzyme activity in cell extracts was assayed as an indication of the ability of the Gal4–VP16C protein to activate transcription. The results shown in Figure 3 are very similar to the results of assaying the mutants in yeast. As in yeast, the mutants with the most significant effects (i.e. activity ≤60%) have substitutions of the phenylalanine residues at positions 473 (49%), 475 (25%) or 479 (22%), or the glutamate residue at 476 (60%). Two other mutants, T458A and H460A, have activities <70% (67 and 62%, respectively) in mammalian cells, although their activities in yeast assays were close to wild type levels. All other alanine substitutions had little or no effect on transcriptional activity in mammalian cells.

Side-chain preferences at key positions in VP16C

The random and systematic mutational analyses described in the preceding sections indicate that F473, F475, E476 and F479 are the residues most critical for the transcriptional function of VP16C in both yeast and mammalian systems. The strong activity of the mutants bearing phenylalanine to leucine substitutions, derived from the PCR-mediated mutagenesis, suggested that the hydrophobic character of the amino acids at those positions is more important than is the aromatic character, in contrast to previous observations for the F442 residue in VP16N. Therefore, to more clearly determine the chemical characteristics most important for the sidechains of the amino acids at these positions, we constructed a number of substitutions at each site and tested the ability of these mutants to activate transcription in yeast and in mammalian cells.

Each of the three key phenylalanine residues was altered to other aromatic amino acids, non-aromatic bulky hydrophobic amino acids, small hydrophobic amino acids, small hydrophilic residues and charged or polar residues. When assayed in yeast, the effects of a particular type of substitution were quite similar at each of the three Phe positions tested (Fig. 4A, B and E). Substitutions of phenylalanine with either aromatic or bulky hydrophobic residues resulted in proteins with at least 50% of wild type activation ability. Even a double mutation to leucine at positions 473 and 475 retained ~60% of wild type activity (Fig. 4B). Alterations at these positions to residues with smaller or hydrophilic sidechains, however, reduced activity to ≤50%. These data support the hypothesis that hydrophobic character is the most significant criterion of these residues that are critical for activation. Interestingly, effects at positions 475 and 479 were more severe than those at 473. Changes resulting in sidechains that are not bulky and hydrophobic only reduced activity to between 50 and 30% for 473, whereas for 475 and 479, activity was <25% for all of these changes. This result may indicate that F473 is less important for transcriptional activation than are F475 and F479.

When assayed in mammalian cells, the substitutions at positions 475 and 479 produced results similar to those obtained in yeast. At these positions, the activities of mutants with hydrophobic amino acid substitutions were at least 60% of wild type respectively, while substitutions with non-hydrophobic residues yielded activities <30% of that of the wild type protein (Fig. 5B and E). Again, these results support the hypothesis that the hydrophobic nature of residues at these positions is important for activation by VP16C. A double mutant with leucines at positions 473 and 475 also retained >30% wild type activity (Fig. 5B).

The indication that F473 is less important than F475 or F479 is demonstrated even more clearly in the mammalian system than in yeast. In the mammalian assays, the only F473 mutant with activity below 75% bore a substitution to alanine (Fig. 5). Thus, whereas F473 is less important in yeast than are the other phenylalanine residues, it contributes very little to the activity of VP16C in the mammalian cell assay.

The only substitution of an acidic residue with alanine that had a significant effect was E476A (Figs 2 and 3). To test the chemical

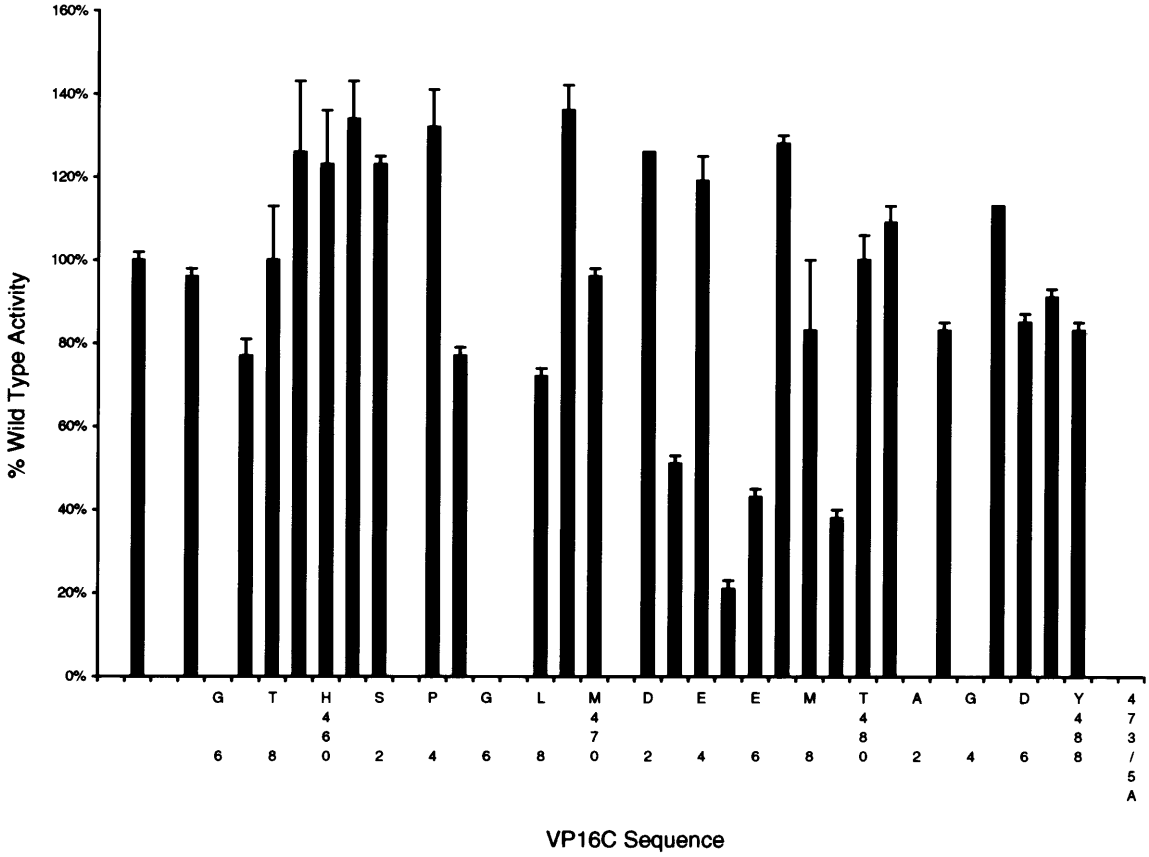


Figure 2. Relative activities in yeast of alanine-scanning mutants of the VP16C subregion. Alanine substitutions were constructed at each position of VP16C except glycines and existing alanines. The activities of Gal4–VP16C fusion proteins bearing these substitutions were assayed using a β-galactosidase reporter gene as described in the legend to Figure 1. The position of each mutation is represented by the amino acid sequence of VP16C along the horizontal axis of this figure. Bars indicate mean activity (with standard deviation) of β-galactosidase in yeast cell extracts from at least three parallel cultures, expressed relative to the activity of the wild type Gal4–VP16C fusion protein (indicated by +). β-galactosidase activities in extracts lacking the Gal4–VP16C fusion protein (indicated by –) were negligible. A double substitution of alanine for both F473 and F475 is shown at the right end of the figure.

characteristics that might be important at that position, mutants were constructed in which the glutamate was replaced with other acidic, uncharged, nonpolar or even basic amino acids (Figs 4C and 5C). The E476D and E476Q mutants functioned with essentially wild type activity in both yeast and mammalian cells, indicating that the charge and length of the sidechain are not critical parameters. Moreover, a mutant with reversed charge (E476K) and mutants with aliphatic rather than polar sidechains (E476M, E476L and E476P) were also at least 50% active in yeast and (with the exception of E476L) also in mammalian cells. The wide range of substitutions that do not significantly affect activity in either assay system suggests that the nature of the side chain at position 476 is not particularly critical for activation by VP16C.

Mutations at position 478 seem to affect the function of VP16C differently in the yeast and mammalian systems. In the yeast assays, leucine was an effective substitute for methionine at this position, retaining 85% activity, while alanine and serine diminished activity to 50 and 40% of wild type levels, respectively (Fig. 4D). In the mammalian assays, however, none of these changes significantly affected activation ability (Fig. 5D). This observation suggests at least some difference in the structures required for activation in the yeast and mammalian systems.

The steady-state levels of all Gal4–VP16C proteins studied in this work were assessed by immunoblot analysis. Expression of Gal4–VP16 mutants in yeast cells was assayed using aliquots of the whole-cell extracts used for β-galactosidase assays. Expression of mutant fusion proteins in mammalian cells was tested by transfecting mouse L cells with 5 μg of each expression plasmid and collecting whole cell extracts 2 days after transfection. Samples containing equivalent amounts of total protein were electrophoresed in SDS–polyacrylamide gels, blotted to nitrocellulose and probed with a polyclonal antiserum raised against recombinant Gal4–VP16 fusion protein. The results of typical immunoblots are shown in Figure 6, which explicitly demonstrates that the mutant proteins with the greatest defects in transcriptional activity are nonetheless present at levels similar to or greater than the wildtype fusion protein. Similar results were obtained for each of the proteins whose transcriptional activities are represented in Figures 2–5 (data not shown). Therefore, the differences in transcriptional activity cannot be attributed to differences in expression or stability of the mutant activator proteins.

DISCUSSION

The first evidence that bulky hydrophobic amino acids, rather than the abundant acidic amino acids, were most critical for the function of an acidic transcriptional activation region was observed

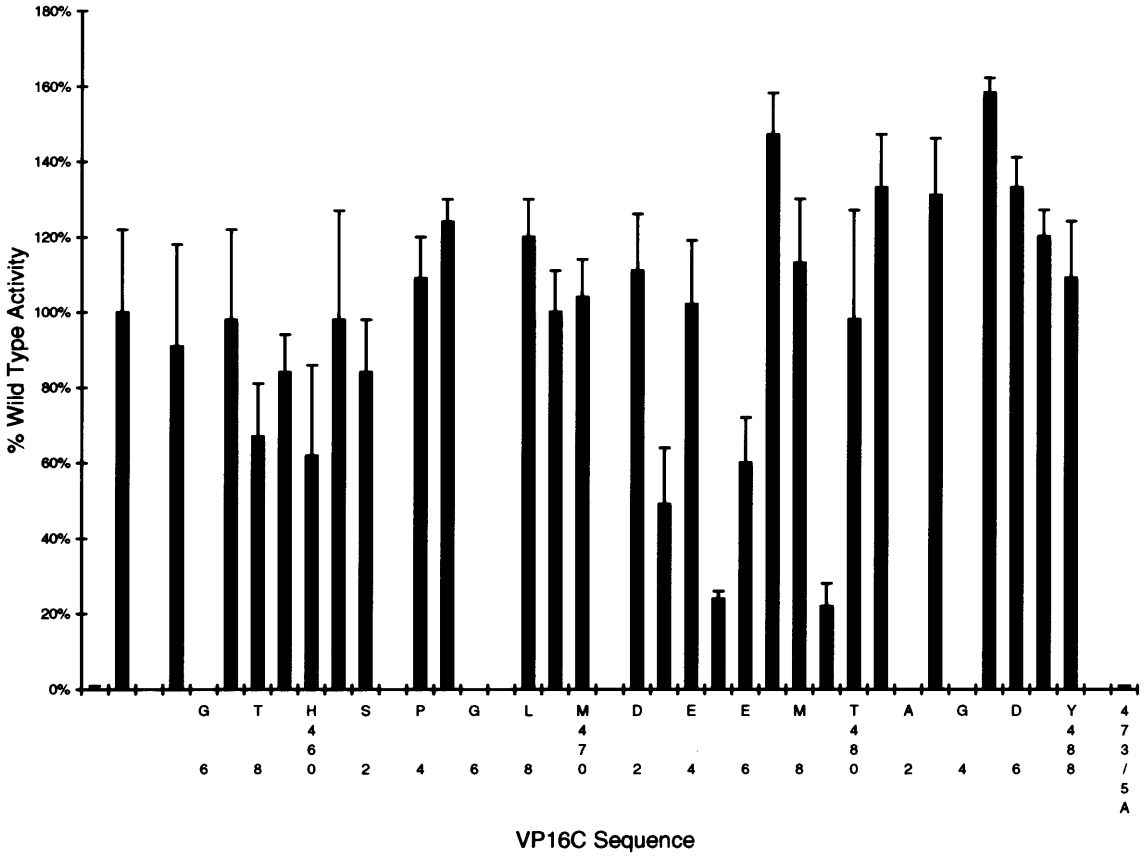


Figure 3. Relative activities of Gal4–VP16C alanine-scanning mutants in mammalian cells. Each of the alanine-scanning mutations described in Figure 2 was recloned into a mammalian expression plasmid and transfected into mouse L cells with a reporter gene that expresses chloramphenicol acetyltransferase under control of Gal4 binding sites (pG_5BCAT). The CAT enzyme activities in cell extracts were assayed by the fluor-diffusion method, adjusted to ensure that measurements were within the linear range of the assay. Bars indicate mean activities (with standard deviations) of CAT activity from at least three plates of cells transfected with a given Gal4–VP16C mutant expression plasmid, relative to the activity of the wild type Gal4–VP16C plasmid (indicated by +).

in studies of the N-subregion of the VP16 transcriptional activation region (VP16N, amino acids 411–456; 7,8). In the present study, we employed two complementary mutational strategies, unbiased with respect to this hypothesis, to identify key amino acids in the VP16 C-subregion (VP16C, amino acids 450–490). The results of both alanine scanning and PCR-mediated random mutagenesis indicated that three phenylalanine residues, at positions 473, 475 and 479, contribute significantly to activation, with F475 being the most critical. Other aromatic or bulky hydrophobic amino acid sidechains apparently contribute relatively little to transcriptional activation by VP16C, because most substitutions replacing Phe457, Tyr465, Leu468, Met470, Ile485 and Tyr488 had little effect on transcriptional activity. We cannot exclude the possibility that amino acid substitutions with more drastic changes in sidechain characteristics might affect transcriptional activity, as suggested by the 60% loss of activity in yeast when Leu468 was replaced by proline (Fig. 1) or when Met478 was altered to serine (Fig. 4D).

Thus, as previously observed for VP16N (7,8), aromatic residues are particularly important for the function of the C-subregion. However, the changes in chemical characteristics of those amino acid sidechains that can be tolerated with little loss of activity differ between the two subregions of VP16. The critical phenylalanine residue of VP16N, F442, could only be functionally replaced by other aromatic residues (8). In contrast, any of the bulky hydrophobic amino acids leucine, isoleucine and valine, as well as cysteine and the aromatic residues tyrosine and tryptophan, could effectively substitute for the key phenylalanines in VP16C (Figs 4 and 5).

Although most of the acidic residues of VP16C could be altered (at least to alanine) without affecting transcriptional activation, our results indicate contributions by the glutamate residues at position 476 and possibly at position 474, two sites flanking the most critical phenylalanine. The alanine scanning results shown in Figures 2 and 3 demonstrate that the single replacement of E476 with alanine reduced the activity of VP16C by half in both yeast and mammalian cells. The negative charge of the amino acid sidechain at this position seems dispensable, however, since mutants bearing non-polar, polar and even basic amino acids at that position retained significant activity (Figs 4C and 5C). No other replacement of an acidic residue with alanine showed any significant effect in our experiments. However, the replacement of E474 with glycine had modest but reproducible effects in the context of the E474G/F475L and F473S/E474G double mutants when compared with the F475L and F473S single mutants (Fig. 1). The transcriptional effect of glycine but not alanine at this position might be due to the destabilizing effects of glycine substitutions on protein secondary structures.

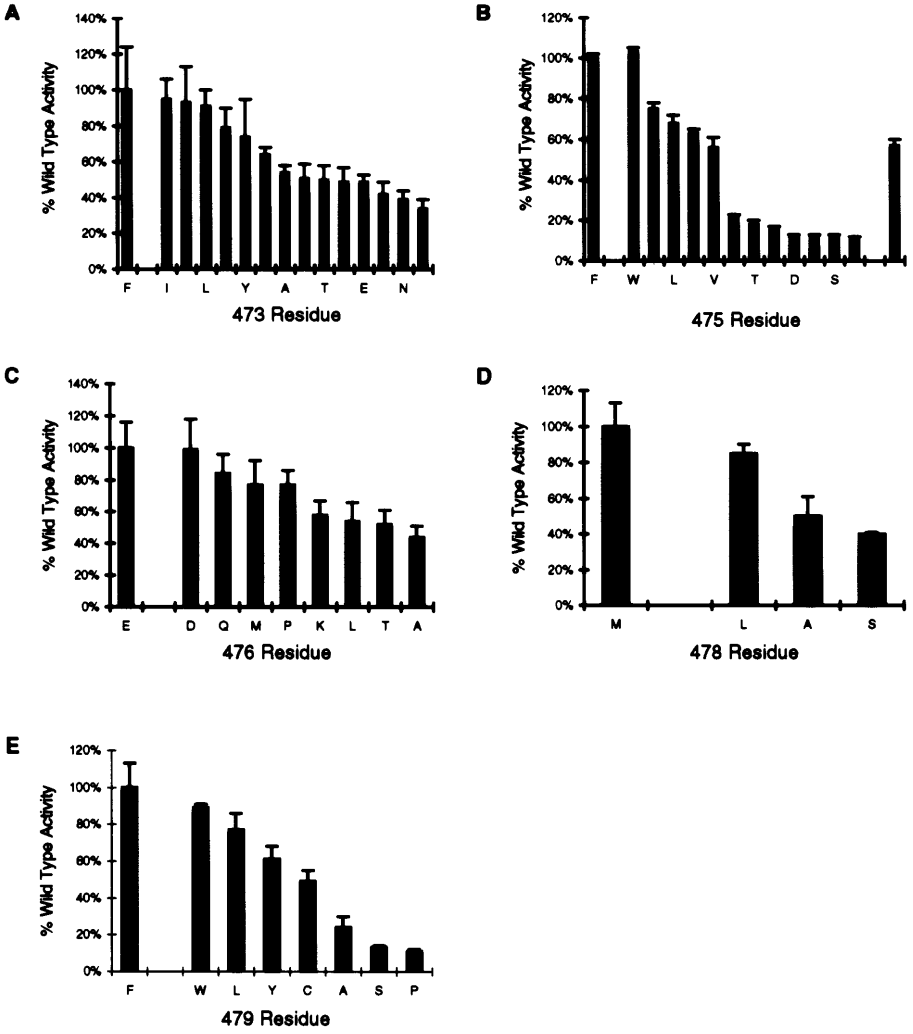


Figure 4. Relative activities in yeast of Gal4–VP16C mutants bearing various amino acid substitutions at specific positions. Mutant Gal4–VP16C proteins with substitutions at F473 (**A**), F475 (**B**), E476 (**C**), M478 (**D**) or F479 (**E**), were assayed for their ability to activate expression of a β-galactosidase reporter gene as described in Figure 1. The individual substitutions are indicated along the horizontal axis of each panel in descending order of activity from left to right, with the mean activity (with standard deviations) from at least three cultures for each mutant being represented by a vertical bar. (B) also indicates the activity of one mutant with leucine substitutions at both F473 and F475.

The relative unimportance of individual acidic amino acids in VP16C is consistent with prior observations of VP16N and several other eukaryotic transcriptional activation regions (7,8,10–15,47,48) in which the negatively charged sidechains seem relatively unimportant for transcriptional activation. For example, when four aspartate residues surrounding the critical Phe442 in VP16N were substituted by asparagine (removing the negative charge) or by glutamate (retaining the charge on a longer sidechain), the asparagine substitutions had less effect (7). Moreover, when single alanine substitutions were tested in VP16N at positions D440, D441, D443 and D445, only the D443A mutation had a significant effect (J.Kastenmayer, P.J.Horn and S.J.Triezenberg, unpublished). Thus, in both VP16C and VP16N the only acidic residue whose replacement by alanine has a significant effect is the residue immediately following the most critical phenylalanine.

In most cases, substitutions in VP16C had roughly equivalent effects when tested in yeast and in mammalian cells. This result implies that similar structural features are required for this activation region to function in cells of these two disparate species, and might be construed to suggest that mechanisms of transcriptional activation are also conserved. An alternative hypothesis is that the specific targets of the activating region differ in yeast and in humans, but that similar structural features of the activator are nonetheless integral to each of those distinct interactions. Careful analysis of protein–protein interactions and the transcriptional consequences of those interactions is necessary to discriminate between these two hypotheses.

Eukaryotic transcriptional activation regions bind *in vitro* to a number of putative target proteins. VP16, for example, has been shown to bind directly to TFIIA (49), TBP (50), dTAF$_{II}$40 or hTAF$_{II}$32 (33,51), TFIIB (52), TFIIH (53) and the yeast adaptor ADA2 (54). Questions remain as to which if any of these interactions are relevant to transcriptional activation *in vivo*, and whether the same hydrophobic interface is employed in each interaction. This question can be addressed in part by detailed

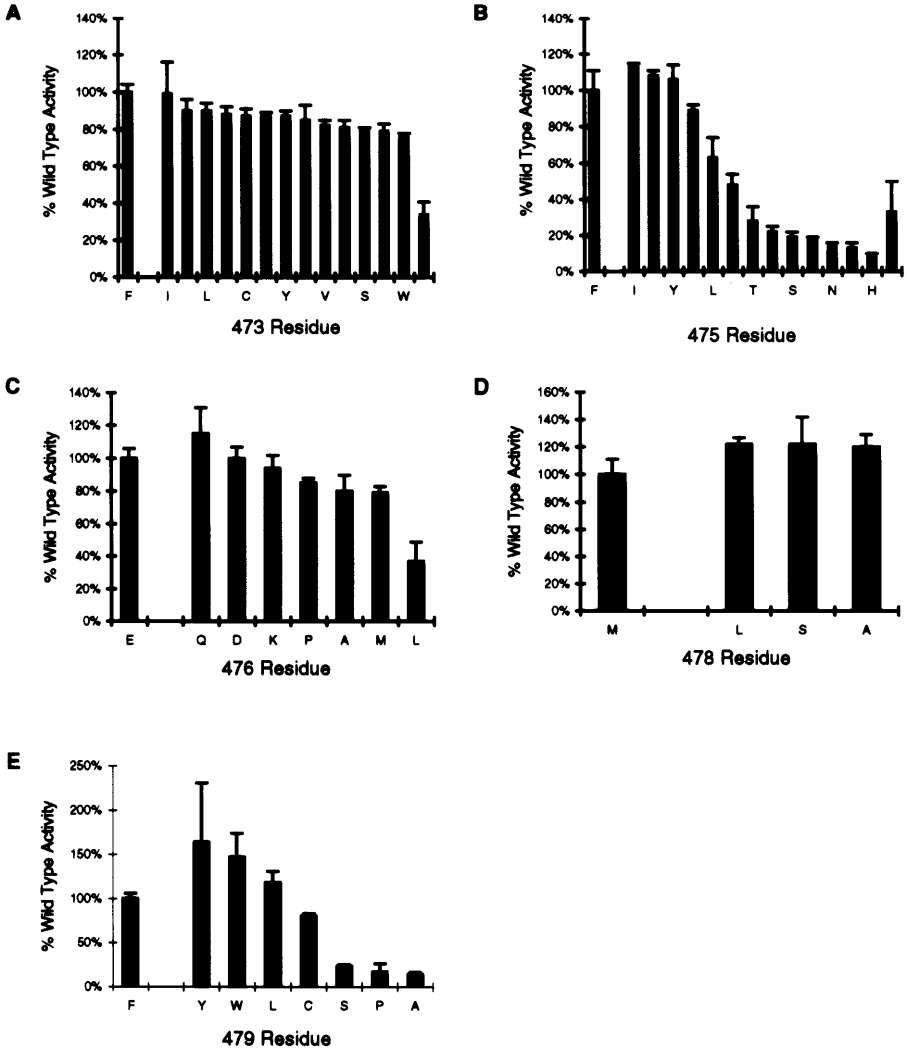


Figure 5. Relative activities in mammalian cells of Gal4–VP16C fusion proteins with various substitutions at specific positions. Each of the VP16C mutants described in Figure 4 was recloned into a mammalian expression vector and assayed for the ability to activate expression of a CAT reporter gene as described in Figure 2. The individual substitutions are indicated along the horizontal axis of each panel in descending order of activity from left to right, with the mean activity (with standard deviations) from at least three plates of transfected cells for each mutant being represented by a vertical bar.

biochemical experiments assessing the roles of activators at particular steps in the transcriptional activation process. For example, the VP16C activation region (and others) can bind to TFIIA and can enhance the rate of formation and the stability of a complex comprising TFIIA, TFIID and promoter DNA (49). When a subset of the VP16C mutants described in this report was tested for this property, a significant correlation was observed between the level of transcriptional activation and the ability to bind TFIIA and to stimulate formation of TFIIA/TFIID/DNA complexes (55). This result is consistent with models invoking recruitment of TFIID as an early step in the mechanism of transcriptional activation (2,56).

Given the propensity of hydrophobic amino acid sidechains to be buried in protein tertiary structures, two simple models can be proposed for the role of such residues in transcriptional activation regions. On one hand, hydrophobic residues may be required for activation because, by interactions with each other, they facilitate formation of a structure that presents other amino acids for interaction with target proteins in the transcriptional machinery. The systematic analysis by alanine-scanning mutagenesis of the VP16C subregion argues against this model, for we found little evidence for the critical role of any other amino acids in transcriptional activation. Alternatively, the critical hydrophobic residues may remain exposed on the surface of the activation region to directly participate in the interaction with other proteins. These residues might then be buried in the interface created when the activation region associates with its target.

The biophysical evidence presently available for VP16 cannot clearly distinguish between these two hypotheses, although results from studies of other activators favor the latter model. Fluorescence spectrosopy of the VP16 activation region revealed that key aromatic residues at positions 442 and 473 were constrained into a less flexible and more hydrophobic environment in the presence of TBP (22), but did not demonstrate that these residues make direct contact with TBP. A recent NMR study revealed that an α helix was induced in VP16C when $hTAF_{II}32$ (a component

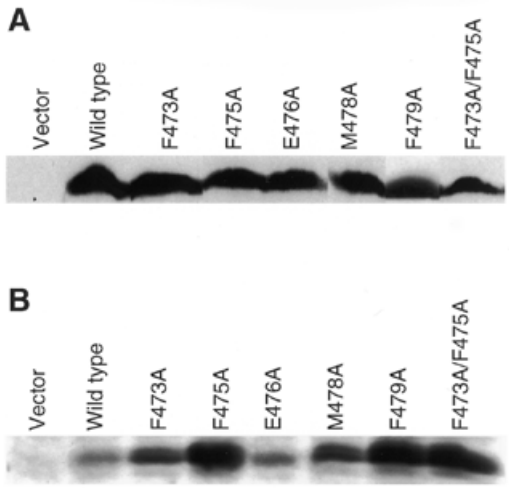


Figure 6. Immunoblot analysis of yeast (**A**) and mouse cells (**B**) expressing wild type or mutant forms of the Gal4–VP16 fusion protein. Aliquots of whole cell extracts were electrophoresed in SDS–polyacrylamide gels and blotted to nitrocellulose. The filters were incubated with polyclonal serum directed against the Gal4–VP16 protein and developed using enhanced chemiluminescence. Extracts from cells transformed with vector only are represented in the left lane of each panel. This representative figure shows the expression levels of mutant proteins with the greatest defects in transcriptional activation. Similar results were obtained for all VP16 mutants tested in both yeast and mammalian cells.

of TFIID) was present (23). Of the three amino acid sidechains within this helix that showed significant changes in chemical shift upon interaction with $TAF_{II}32$, two (F479 and L483) are bulky hydrophobic residues and the third (D472) is acidic. However, these changes in chemical shift do not necessarily imply that those sidechains directly participate in the interaction. Moreover, the results of this NMR analysis are at odds with the mutational analysis described in the present report, in which F473 and especially F475 are shown to play key roles in transcriptional activation whereas alanine substitutions of D472 and L483 had no apparent effect. Thus, one might have expected the NMR experiments to show pronounced chemical shifts involving the sidechains of F473 and F475. It may be that these residues are not involved in interactions with $hTAF_{II}32$, but are critical for some other interaction required for transcriptional activation. Alternatively, the *in vitro* interaction of VP16C with $hTAF_{II}32$ may not represent an essential step in the process of transcriptional activation. We are presently testing our panel of VP16C mutations at these positions for their effects on binding to $hTAF_{II}32$ and other potential target proteins (P.J.Horn, Y.Nedialkov and S.J.Triezenberg, unpublished results).

Two studies offer strong evidence that the hydrophobic residues of some activating regions form surfaces that interact directly with target proteins. In a crystallographic analysis of a peptide representing the p53 transcriptional activation region complexed with its repressor MDM2 (24), the p53 peptide folded into a helix in which transcriptionally-important hydrophobic residues formed the interacting surface. The NMR solution structure of a complex containing a peptide from the activation region of CREB and its interacting region from the coactivator protein CBP also showed a transition of the activation region from random coil to amphipathic helix (25). In this case, the interaction surface involved hydrophobic residues of the activation region and a hydrophobic groove of the coactivator, reinforced by electrostatic and hydrogen-binding interactions contributed by acidic and phosphoserine residues of the activator. We hypothesize that a similar hydrophobic surface of VP16C, encompassing the phenylalanine residues at 475 and 479, may interact with target proteins at key steps along the pathway to transcriptional activation, but the identification of those key steps, the relevant target proteins at those steps, and the structure of VP16 when interacting with those target proteins, await further study.

The mutational, biochemical and biophysical studies summarized above lead to a model in which transcriptional activation by VP16 and other eukaryotic activators requires biochemical interactions between activation regions and specific target proteins that depend most directly upon hydrophobic residues within helical segments of the activator, with relatively minor or non-specific contributions made by acidic sidechains. The biophysical evidence that in at least some circumstances these hydrophobic residues are present in helical segments induced by the presence of the target protein recalls to mind the amphipathic helix model introduced early in the study of eukaryotic activators (57,58). This model originally posited that activation regions would fold into amphipathic α-helices, the acidic face of which would comprise the interaction surface with targets such as the basic regions of TBP and TFIIB. The present evidence, while largely consistent with the formation of helices in activating regions, suggests instead that hydrophobic surfaces are key to interactions with target proteins, whereas the acidic residues may be present for bringing these otherwise buried regions to the surface, for long-range electrostatic interactions that might facilitate docking, or for minor supporting roles in the induced interaction structure. A remaining question is whether the same secondary structure and hydrophobic interaction surface will suffice for interactions of a given activator, such as VP16, with each of its various putative target proteins, and if so how this promiscuous interaction surface can nonetheless retain sufficient specificity to effectively accomplish its transcriptional task in the complex environment of the eukaryotic nucleus.

ACKNOWLEDGEMENTS

The aim of this study was originally defined by S.J.T. and A.K. The alanine-scanning strategy was proposed by R.H.E. and the mutants initially constructed by W.N. P.J.H. isolated and characterized the PCR-generated mutants. The alanine scanning and other site-directed mutants were tested by S.M.S. with help from V.A.O. and A.H.K. The authors thank Rick Babbitt for DNA preparation and Lee Alexander for antibody production and technical support. This work was supported by NIH grants AI27323 and AI01284 to S.J.T. and GM51107 to A.H.K. R.J.E. is supported in part by an HHMI Investigatorship.

REFERENCES

1 Kingston,R.E., Bunker,C.A. and Imbalzano,A.N. (1996) *Genes Dev.*, **10**, 905–920.
2 Ptashne,M. and Gann,A. (1997) *Nature*, **386**, 569–577.
3 Triezenberg,S.J. (1995) *Curr. Opin. Genet. Dev.*, **5**, 190–196.
4 Bentley,D.L. (1995) *Curr. Opin. Genet. Dev.*, **5**, 210–216.
5 Mitchell,P.J. and Tjian,R. (1989) *Science*, **245**, 371–378.
6 Attardi,L.D. and Tjian,R. (1993) *Genes Dev.*, **7**, 1341–1353.
7 Cress,W.D. and Triezenberg,S.J. (1991) *Science*, **251**, 87–90.
8 Regier,J.L., Shen,F. and Triezenberg,S.J. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 883–887.

9 Blair,W.S., Bogerd,H.P., Madore,S.J. and Cullen,B.R. (1994) *Mol. Cell. Biol.*, **14**, 7226–7234.
10 Lin,J., Chen,J., Elenbaas,B. and Levine,A.J. (1994) *Genes Dev.*, **8**, 1235–1246.
11 Lee,S.W., Chang,J., Lee,C.W., Kim,D.H., Choi,K.Y. and Sung,Y.C. (1995) *Mol. Cells*, **5**, 467–474.
12 Jackson,B.M., Drysdale,C.M., Natarajan,K. and Hinnebusch,A.G. (1996) *Mol. Cell. Biol*, **16**, 5557–5571.
13 Iniguez-Lluhi,J.A., Lou,D.Y. and Yamamoto,K.R. (1997) *J. Biol. Chem.*, **272**, 4149–4156.
14 Almlöf,T., Gustafsson,J.Å. and Wright,A.P.H. (1997) *Mol. Cell. Biol.*, **17**, 934–945.
15 Sainz,M.B., Goff,S.A. and Chandler,V.L. (1997) *Mol. Cell. Biol.*, **17**, 115–122.
16 Gill,G., Pascal,E., Tseng,Z.H. and Tjian,R. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 192–196.
17 Donaldson,L. and Capone,J.P. (1992) *J. Biol. Chem.*, **267**, 1411–1414.
18 O'Hare,P. and Williams,G. (1992) *Biochemistry*, **31**, 4150–4156.
19 Schmitz,M.L., dos Santos Silva,M.A., Altmann,H., Czisch,M., Holak,T.A. and Baeuerle,P.A. (1994) *J. Biol. Chem.*, **269**, 25613–25620.
20 Van Hoy,M., Hansen,A. and Kodadek,T. (1992) *J. Am. Chem. Soc.*, **114**, 362–363.
21 Shen,F., Triezenberg,S.J., Hensley,P., Porter,D. and Knutson,J.R. (1996) *J. Biol. Chem.*, **271**, 4819–4826.
22 Shen,F., Triezenberg,S.J., Hensley,P., Porter,D. and Knutson,J.R. (1996) *J. Biol. Chem.*, **271**, 4827–4837.
23 Uesugi,M., Nyanguile,O., Lu,H., Levine,A.J. and Verdine,G.L. (1997) *Science*, **277**, 1310–1313.
24 Kussie,P.H., Gorina,S., Marechal,V., Elenbaas,B., Moreau,J., Levine,A.J. and Pavletich,N.P. (1996) *Science*, **274**, 948–953.
25 Radhakrishnan,I., Perez-Alvarado,G.C., Parker,D., Dyson,H.J., Montminy,M.R. and Wright,P.E. (1997) *Cell*, **91**, 741–752.
26 O'Hare,P. (1993) *Semin. Virol.*, **4**, 145–155.
27 Triezenberg,S.J., Kingsbury,R.C. and McKnight,S.L. (1988) *Genes Dev.*, **2**, 718–729.
28 Hayes,S. and O'Hare,P. (1993) *J. Virol.*, **67**, 852–862.
29 Sadowski,I., Ma,J., Triezenberg,S. and Ptashne,M. (1988) *Nature*, **335**, 563–564.
30 Cousens,D.J., Greaves,R., Goding,C.R. and O'Hare,P. (1989) *EMBO J.*, **8**, 2337–2342.
31 Seipel,K., Georgiev,O. and Schaffner,W. (1992) *EMBO J.*, **11**, 4961–4968.
32 Walker,S., Greaves,R. and O'Hare,P. (1993) *Mol. Cell. Biol.*, **13**, 5233–5244.
33 Goodrich,J.A., Hoey,T., Thut,C.J., Admon,A. and Tjian,R. (1993) *Cell*, **75**, 519–530.
34 Cunningham,B.C. and Wells,J.A. (1989) *Science*, **244**, 1081–1085.
35 Guarente,L., Yocum,R.R. and Gifford,P. (1982) *Proc. Natl Acad. Sci. USA*, **79**, 7410–7414.
36 Lillie,J.W. and Green,M.R. (1989) *Nature*, **338**, 39–44.
37 Leung,D.W., Chen,E. and Goeddel,D.V. (1989) *Technique*, **1**, 11–15.
38 Spee,J.H., deVos,W.M. and Kuipers,O.P. (1993) *Nucleic Acids Res.*, **21**, 777–778.
39 Berger,S.L., Pina,B., Silverman,N., Marcus,G.A., Agapite,J., Regier,J.L., Triezenberg,S.J. and Guarente,L. (1992) *Cell*, **70**, 251–265.
40 Kunkel,T.A. (1985) *Proc. Natl Acad. Sci. USA*, **82**, 488–492.
41 Gietz,R.D., St. Jean,A., Woods,R.A. and Schiestl,R.H. (1992) *Nucleic Acids Res.*, **20**, 1425–1427.
42 Rose,M. and Botstein,D. (1983) *Methods Enzymol.*, **101**, 167–180.
43 Bradford,M.M. (1976) *Anal. Biochem.*, **72**, 248–254.
44 Lopata,M.A., Cleveland,D.W. and Sollner-Webb,B. (1984) *Nucleic Acids Res.*, **12**, 5707–5717.
45 Nielsen,D.A., Chang,T. and Shapiro,D.J. (1989) *Anal. Biochem.*, **179**, 12–23.
46 Neumann,J.R., Morency,C.A. and Russian,K.O. (1987) *BioTechniques*, **5**, 444–447.
47 John,M., Briand,J.-P. and Schnarr,M. (1996) *Peptide Res.*, **9**, 71–78.
48 Blair,W.S., Bogerd,H. and Cullen,B.R. (1994) *J. Virol.*, **68**, 3803–3808.
49 Kobayashi,N., Boyer,T.G. and Berk,A.J. (1995) *Mol. Cell. Biol.*, **15**, 6465–6473.
50 Ingles,C.J., Shales,M., Cress,W.D., Triezenberg,S.J. and Greenblatt,J. (1991) *Nature*, **351**, 588–590.
51 Klemm,R.D., Goodrich,J.A., Zhou,S.L. and Tjian,R. (1995) *Proc. Natl Acad. Sci. USA*, **92**, 5788–5792.
52 Lin,Y.S., Ha,I., Maldonado,E., Reinberg,D. and Green,M.R. (1991) *Nature*, **353**, 569–571.
53 Xiao,H., Pearson,A., Coulombe,B., Truant,R., Zhang,S., Regier,J.L., Triezenberg,S.J., Reinberg,D., Flores,O., Ingles,C.J. and Greenblatt,J. (1994) *Mol. Cell. Biol.*, **14**, 7013–7024.
54 Barlev,N.A., Candau,R., Wang,L., Darpino,P., Silverman,N. and Berger,S.B. (1995) *J. Biol. Chem.*, **270**, 19337–19344.
55 Kobayashi,N., Horn,P.J., Sullivan,S.M., Triezenberg,S.J., Boyer,T.G. and Berk,A.J. (1998) *Mol. Cell. Biol.*, **18**, 4023–4031.
56 Stargell,L.A. and Struhl,K. (1996) *Trends Genet.*, **12**, 311–315.
57 Giniger,E. and Ptashne,M. (1987) *Nature*, **330**, 670–672.
58 Ptashne,M. (1988) *Nature*, **335**, 683–689.