

Strain Variations in Single Amino Acids of the 86-Kilodalton Human Cytomegalovirus Major Immediate-Early Protein (IE2) Affect Its Functional and Biochemical Properties: Implications of Dynamic Protein Conformation

M. Inmaculada Barrasa, Noam Harel, Yongjun Yu, and James C. Alwine*

Department of Cancer Biology, Abramson Family Cancer Research Institute, School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6142

Received 22 August 2002/Accepted 19 December 2002

The 86-kDa major immediate-early protein, IEP86 (IE2, IE2_{579aa}, or ppUL122a), from the Towne and AD169 strains of human cytomegalovirus show four amino acid variations, namely, R68Q, K455E, T541A, and seven consecutive serines beginning at position 258 in Towne and eight serines in AD169. A commonly utilized IEP86 cDNA expression clone (herein called the original cDNA) (E. Baracchini, E. Glezer, K. Fish, R. M. Stenberg, J. A. Nelson, and P. Ghazal, *Virology* 188:518–529, 1992) shows the Towne R68 and seven serines but contains the AD169 E455 and A541 plus two amino acid mutations, M242I and A463T. In transcriptional activation analyses using several promoters, the IEP86 produced by the original cDNA was 40 to 60% less active than wild-type (WT) Towne IEP86, whereas AD169 IEP86 was two to three times more active than WT Towne IEP86. To determine which amino acid variations or mutations accounted for the differences in transcriptional activation, they were individually tested in the WT Towne IEP86 background. K455E, M242I, and the eighth serine had little effect on transcriptional activation or sumoylation when inserted into the Towne background. T541A significantly increased transcriptional activation on all promoters tested and showed increased sumoylation; T541A is the primary reason that WT AD169 IEP86 has increased activity over WT Towne IEP86. The increased sumoylation seen with T541A was quantitatively reduced to WT Towne levels when the K455E alteration was present, suggesting that K455 may be a sumoylation site or that E455 may cause alterations in the IEP86 structure which affect overall sumoylation. A463T was very deleterious to transcriptional activation and caused reduced sumoylation. The A436T mutation in the original cDNA is partially compensated by the presence of the T541A variation. Phosphopeptide mapping suggests that a threonine at 463 or 541 does not introduce a phosphorylation site. However, the A463T mutation does affect phosphorylation at a distant site, suggesting that it alters the conformation of the protein. Promoter-specific effects were noted with some of the amino acid variations, particularly T541A. Structural modeling is presented which suggests how A463T and T541A alter the functional structure of WT Towne IEP86. A hydrophobic core containing A463 is predicted to be responsible for the functional integrity of the carboxy-terminal region of IEP86 between amino acids 344 and 579.

Transcription of the major immediate-early gene of human cytomegalovirus (HCMV) results in a primary transcript which can be alternatively spliced and polyadenylated to form several mRNAs encoding different major immediate-early proteins (MIEPs) (26, 32, 33, 35). Two of these proteins, IEP72 (72 kDa; also called IE1, IE1_{491aa}, or ppUL123) and IEP86 (86 kDa; also called IE2, IE2_{579aa}, or ppUL122a), appear in greater abundance in the lytic infection and have been extensively examined because many studies have shown that they affect RNA polymerase II (pol II) transcription in a cell- and promoter-specific manner. These effects alter the transcriptional activity of viral and cellular promoters and provide the control of the temporal expression of the viral early and late genes (6, 10, 11, 14, 16, 19, 20, 23–25, 27, 30, 31, 34, 39, 41). The MIEPs have also been shown to affect cell cycle control and apoptosis (22, 40, 42).

Many studies of the MIEPs use two highly passaged HCMV strains, AD169 and Towne, both of which were originally developed for vaccine use (18, 29). Comparison of the sequences for IEP86 from each of the strains shows that there are four amino acid variations between Towne and AD169. These include the three amino acid substitutions R68Q, K455E, and T541A (amino acid numbering is that of the wild-type [WT] Towne IEP86) and the insertion of a serine within a cluster of serines between 258 and 264; there are seven consecutive serines in Towne and eight in AD169 (Fig. 1). In the following work, we show that some of these variations affect IEP86 activity and function.

In some of our studies of IEP86, we have used an IEP86 cDNA expression plasmid (3), herein called the IEP86 original cDNA (OcDNA). Our sequence analysis shows that although the OcDNA contains the Towne strain's R68 and seven serines beginning at position 258, the amino acids at positions 455 and 541 were those of AD169 (E455 and A541). In addition, there were two amino acid substitutions in the cDNA which are not seen in either Towne or AD169, namely, M242I and A463T

* Corresponding author. Mailing address: 314 Biomedical Research Building, 421 Curie Blvd., University of Pennsylvania, Philadelphia, PA 19104-6142. Phone (215) 898-3256. Fax: (215) 573-3888. E-mail: alwine@mail.med.upenn.edu.

| Sequences | 68 | 242 | 455* | 463* | 541* | # Serines at 258* |
|---------------------|----|-----|------|------|------|-------------------|
| IEP86 Towne | R | M | K | A | T | 7 |
| IEP86 AD169 | Q | M | E | A | A | 8 |
| IEP86 Original cDNA | R | I | E | T | A | 7 |
| cDNAs made: | | | | | | |
| IEP86 Towne cDNA | R | M | K | A | T | 7 |
| T541A | R | M | K | A | A | 7 |
| A463T | R | M | K | T | T | 7 |
| K455E | R | M | E | A | T | 7 |
| M242I | R | I | K | A | T | 7 |
| +S258 | R | M | K | A | T | 8 |
| K455E/T541A | R | M | E | A | A | 7 |
| K455E/T541A +S258 | R | M | E | A | A | 8 |
| A463T/T541A | R | M | K | T | A | 7 |
| K455E/A463T | R | M | E | T | T | 7 |
| IEP86 AD169 cDNA | Q | M | E | A | A | 8 |

*Towne AA numbering;
AD169 is +1 after 258.

FIG. 1. IEP86 variants and mutants used in the studies presented. The top segment shows the amino acid variations and mutations between Towne IEP86 and AD169 IEP86 and the IEP86 encoded by the OcDNA. The bottom segment shows the individual variants and mutants made for the present studies.

(Fig. 1). The A463T mutation and its effects have been observed by others (Richard Greaves, personal communication). The source of these mutations is not known.

In the studies presented, we found that the IEP86 produced by the OcDNA was 40 to 60% less active than WT Towne IEP86 in transcriptional activation. However, promoter specificity was indicated; the most significant reduction was seen with a synthetic, nonviral promoter (Tef-TATA) and an HCMV late promoter (ICP36), whereas an HCMV early promoter (UL112-113) was less affected by the variations and mutations, suggesting that the amino acid variations and mutations affect some promoters more significantly than others. In contrast, WT AD169 IEP86 was found to activate all the promoters two- to threefold more than WT Towne IEP86. To determine which amino acids accounted for these effects, we introduced the amino acid variations and mutations individually (and in combination) into the WT Towne IEP86 background (Fig. 1). These were compared to WT Towne IEP86 with respect to (i) transcriptional activation, (ii) sumoylation, which has been shown to affect transcriptional activation (1, 15), and (iii) phosphorylation. We found that the K455E variation and the M242I mutation had little effect on transcriptional activation and sumoylation. Adding the eighth serine at 258 also had only modest effects but indicated promoter specificity by preferentially activating the UL112-113 promoter. However, the amino acid variation T541A increased transcriptional activation on all three promoters by as much as threefold compared to WT Towne IEP86; in addition, the T541A IEP86 showed increased sumoylation in the Towne background. However, this increased sumoylation was modulated to WT Towne levels by the addition of K455E. Thus, WT Towne and WT AD169 IEP86 are similar, quantitatively, in sumoylation. However, the sites of sumoylation may be different, since K455 could be a sumoylation site or E455 could alter the protein conformation such that overall sumoylation is affected.

Introduction of the A463T mutation into the Towne background was very deleterious to transcriptional activation of all

promoters and caused reduced sumoylation. The ability of the OcDNA IEP86 to maintained transcriptional activation despite the presence of the A463T mutation was shown to be primarily due to the presence of the T541A variation. In addition, the T541A alteration is the primary reason why WT AD169 IEP86 was a better transcriptional activator than WT Towne IEP86. The transcriptional activation results, in combination with phosphopeptide analysis and the sumoylation studies, suggested that the T541A variation and the A463T mutation affect the functional structure of IEP86. Structural modeling of the WT Towne IEP86 is presented which predicts how A463T and T541A affect structure. A463 is predicted to be in a hydrophobic core important for the integrity of the structure of the carboxy-terminal region of IEP86 between amino acids 344 and 579.

MATERIALS AND METHODS

Plasmids. The expression plasmids used included the following. pRL43a contains the genomic major immediate early (MIE) region from the Towne strain (28) and expresses all of the MIEPs. pIE86 (34) and pRSV86 (3, 9) have been described previously. Each contains the same cDNA encoding IEP86 under the control of either the HCMV MIE promoter (pIE86) or the Rous sarcoma virus long terminal repeat (pRSV86). Sequencing of pRSV86 indicated the variations from the Towne sequence as described above and shown in Fig. 1. A fully Towne IEP86-expressing plasmid (pIE86Towne) was made by taking the *SmaI-Bsu36I* fragment from within exon 7 of the MIE gene (taken from pRL43a) and inserting it to replace the same fragment in the original pIE86. This fragment spans amino acids 139 to 573 of IEP86 and contains all the amino acids which needed to be corrected or converted from AD169 to Towne. A *HpaI-HindIII* insert from pIE86Towne was transferred to the original pRSV86 to produce pRSV86Towne. Single amino acid changes R68Q, M242I, +S258, K455E, A463T, and T541A were introduced in pRSV86Towne by using the QuikChange site-directed mutagenesis kit (Stratagene) with the oligonucleotides R68Q (5' oligonucleotide, 5'ctttgaaacaagtgaccg3'; 3' oligonucleotide, 5'cggtcactctgtcaaaag3'), M242I (5' oligonucleotide, 5'ctctccagataaacaccctc3'; 3' oligonucleotide, 5'gggtggttatctggagagcgg3'), K455E (5' oligonucleotide, 5'gaagtggcgcagcgca3'; 3' oligonucleotide, 5'tgcgctgcgccactctgggtggatgtgtcac3'), A463T (5' oligonucleotide, 5'actgtaacgaaggcgtc3'; 3' oligonucleotide, 5'gccctctgtacaggtatcggtgcgctg3'), T541A (5' oligonucleotide, 5'gcaaggcctacgcccgtg3'; 3' oligonucleotide, 5'ggcgtaggcctcgcgcccgtctctagat3'), and +S258 (5' oligonucleotide, 5'ccggcccgatgaagatagcagctctctctctc3'; 3' oligonucleotide, 5'ggaagacgaagaggaactgctatctcctcggccg3'). Multiple amino acid alterations (Fig. 1) were made by combining the above by successive site-directed mutagenesis or restriction fragment swapping. All mutations were verified by sequence analysis.

The reporter plasmids used included the following. Dmp7-LUC, a luciferase reporter, contains a simple synthetic promoter consisting of six copies of the Tef-1 element, derived from the simian virus 40 (SV40) late promoter, upstream of the β -globin TATA element (37); this promoter has previously been shown to be activated efficiently by the MIEPs (24). pHM142 contains the luciferase reporter gene under the control of the promoter for the HCMV UL112-113 early genes (2). pICP36 is a luciferase reporter using the promoter of the HCMV late gene ICP36 (UL44) (21). We also utilized pHM976 (a gift of Thomas Stammers), which directs the expression of the fusion protein FLAG-SUMO-1 (15).

Cells, transfections, and infections. The glioblastoma-astrocytoma cell line U-373MG was maintained at passage numbers less than 30. Cells were cultured in Dulbecco's modification of high-glucose Eagle's medium supplemented with 10% fetal calf serum, glutamax, and antibiotics. Cells were transfected with Fugene (Roche) by following the manufacturer's instructions. For luciferase assays, 3×10^5 cells were plated per well in 12-well plates. A transfection mix was made containing 0.2 μ g of the enhanced green fluorescent protein-expressing plasmid pCMS-EGFP (Clontech), 0.5 μ g of reporter plasmid, and 0.1, 0.5, or 1.2 μ g of IEP86-expressing plasmids plus control plasmid pRSV3/BgIII (8) to make up to 1.2 μ g. Four wells of the 12-well plate were transfected with one-quarter of these transfection mixtures each. Three of these wells were used for luciferase assays, and the fourth was used to obtain total extract to assay for IEP86 protein expression. Luciferase activity was assayed by using the Promega luciferase assay system with 12 μ g of transfected cell extract harvested in accordance with the

manufacturer's instructions and was measured on a Berthold 9501 Lumat luminometer.

For studies of sumoylation in infected cultures, 8×10^5 U-373MG cells were grown on 60-mm-diameter tissue culture plates and infected with WT Towne or WT AD169 strains of HCMV at multiplicities of infection (MOIs) of 1 and 4. Twenty-four hours after being infected, the cultures were harvested and analyzed as described below.

Antibodies. Monoclonal antibody MAB810, which recognizes the N-terminal region of IEP2, was purchased from Chemicon and used to immunoprecipitate IEP86. For Western analysis, IEP86 was detected with anti-exon 2/3, a rabbit polyclonal antibody made against the amino acids encoded in exon 2/3 of the MIE gene fused to glutathione *S*-transferase. These amino acids make up the common amino acids shared among the known MIEPs. The anti-FLAG M2 monoclonal antibody was purchased from Upstate Biotechnology. The exon 5 (IEP86)-specific polyclonal antibody anti-pHM178 used to detect IEP86 and its sumoylated forms was the gift of Thomas Stamminger.

Immunoprecipitation and Western blot analysis. Approximately 1.6×10^6 cells were plated on 60-mm-diameter plates. Cells were transfected with 1.8 μ g of the IEP86-expressing plasmids plus 0.2 μ g of pCMS-EGFP. Total cell extracts were made by lysing the cells in RIPA buffer (1% NP-40, 1% deoxycholate, 0.1% sodium dodecyl sulfate [SDS], 150 mM NaCl, 10 mM sodium phosphate [pH 7.2], 2 mM EDTA) plus 1 μ g of leupeptin/ml, 0.7 μ g of pepstatin/ml, 1 mM phenylmethylsulfonyl fluoride, 10 μ g of aprotinin/ml, 2.5 μ g of E64 [trans-Epoxysuccinyl-L-leucylamido-(4-guanidino)butane], 4 mM NaF, and 0.1 mM sodium orthovanadate. The lysate was centrifuged for 30 min at 4°C. The supernatants were used for immunoprecipitation by overnight treatment of 1 mg of extract, 1 μ l of MAB810, and 45 μ l of protein G agarose beads (75% slurry) at 4°C. The samples were then washed, eluted, and separated by SDS-8% polyacrylamide gel electrophoresis. The gel was transferred to nitrocellulose and probed with anti-exon 2/3, followed by incubation with the anti-rabbit secondary antibody. The blot was developed with enhanced chemiluminescence (Amersham).

For the studies of sumoylation, 60-mm-diameter plates were cotransfected with 0.9 μ g of the IEP86-expressing plasmid, 0.9 μ g of pHM976, the FLAG-SUMO-expressing plasmid, and 0.2 μ g of plasmid pCMS-EGFP. IEP86 proteins were immunoprecipitated from 3 mg of total extract, 1 μ l of MAB810 (Chemicon), and 45 μ l of protein G agarose beads (75% slurry) overnight at 4°C. Sumoylated forms of IEP86 were detected by Western analysis as described above by probing with anti-FLAG M2 monoclonal antibody.

Phosphopeptide analysis. Cell labeling, harvest, and tryptic phosphopeptide mapping was performed as previously described (12). Briefly, U373MG cells were transfected with plasmids expressing WT or alternate forms of IEP86. The cells were labeled with [³²P]orthophosphate, and the IEP86 was immunoprecipitated from cell extracts, separated on a SDS-9% polyacrylamide gel electrophoresis gel, excised from gel slices, and digested with trypsin; the phosphopeptide map was determined by electrophoresis and thin-layer chromatography (4).

Secondary structure predictions. IEP86 secondary structure was predicted by comparing the estimations made with the programs SAM-T99 (<http://www.cse.ucsc.edu/research/compbio/HMM-apps/T99-query.html>), PSIPred (<http://bioinf.cs.ucl.ac.uk/psipred/>), PROFsec and PHD (<http://cubic.bioc.columbia.edu/predictprotein/>), and Jpred (<http://jura.ebi.ac.uk:8888/>). Based on the evaluation of automatic protein structure prediction (EVA) scores (<http://maple.bioc.columbia.edu/eva/index.html>) and the critical assessment of fully automated structure prediction-2 (CAFASP-2) evaluation results (<http://www.cs.bgu.ac.il/~dfischer/CAFASP2/eval.html>), SAM-T99, PSIPred, and PROFsec were selected for the prediction of the C-terminal fragment of IEP86 (amino acids 344 to 579). These programs statically gave the most reliable results, with results of 74 to 78% for Q3 (the percentage of residues predicted correctly for the three conformational states, strand, helix, and loop) on the common subset predictions.

The SAM-T99 and PSIPred programs used a single sequence (HCMV IEP86) as the input and looked for proteins related to the query by using the hidden Markov model and PSI-BLAST, respectively. Predictions were made based on multiple sequence alignment (MSA). For the PROFsec prediction of the C-terminal end (amino acids 344 to 579), we used a MSA made with PILEUP (GCC program suite, Wisconsin Package version 10.3; Accelrys Inc., San Diego, Calif.); the sequences included HCMV Towne IEP86 and the homologous proteins of simian, *Macaca mulatta*, tupaia (tree shrew), rat, and murine cytomegaloviruses as well as human herpesvirus 6 (HHV-6) IE2 and HHV-7 IE-A. Adding HHV-6 and HHV-7 proteins to the alignment significantly increased the confidence of the predictions.

Tertiary structure predictions. The C-terminal fragment (amino acids 344 to 579), which includes the dimerization domain and most of the DNA binding domain (amino acids 330 to 579), was modeled by using the three-dimensional

prediction program THREADER, version 2.5 (17). The best fitting model was one based on the structure of *Bacillus thuringiensis* delta-endotoxin cryIA(A). Homology modeling was carried out with Insight II (Accelrys, Inc.) and was followed by energy minimization and molecular dynamics calculations that utilized the CHARMM module and force field (5) to eliminate inappropriate structures.

RESULTS

The AD169 and Towne IEP86 sequences. The literature and databases contain various GenBank entries for the IEP86 sequence of AD169 and Towne. The AD169 IEP86 sequence we used was derived from the genomic sequence by joining the complement of nucleotides 169364 to 170850 (exon 5), 172396 to 172580 (exon 3), and 172695 to 172765 (exon 2) (accession number NC_001347). The first 85 amino acids of Towne IEP86 were derived from the sequence of Towne IEP72 (IE1; accession number AAA45979), since the exons encoding these amino acids are shared with IEP86. The rest of the sequence, unique to IEP86, is encoded by exon 5/7. There are two nucleotide entries for Towne exon 5/7, M11298 (36) and M26973 (33). The two sequences differ in several places. In particular, the regions encoding amino acids 111 to 116 are FRSVRR (ttc cgc agc gtt aga cgc) in entry M26973 and SCSVSS (tct tgc agc gtt agc agc) in entry M11298. Our sequencing of pRV86 (OcDNA) and pRL43a, a plasmid encoding the genomic Towne MIE gene (28), showed that each has the sequence of entry M11298; thus, we conclude that the amino acid sequence from positions 111 to 116 is SCSVSS for WT Towne IEP86. In agreement, this is also the sequence found in AD169. Other differences were resolved by comparison to the Towne genomic sequences in pRL43a.

Transcriptional activation by Towne IEP86, AD169 IEP86, and the OcDNA IEP86. We first corrected all of the amino acid differences in the original cDNA to those of the Towne strain and to those of the AD169 strain. We then compared the ability of each form of IEP86 to transcriptionally activate a number of promoters in transient transfection experiments. The luciferase reporter plasmids contained the following promoters: (i) the Tef-TATA promoter, which was previously shown to be well activated by the MIEPs (24) and which is a synthetic promoter containing the TATA element from the β -globin gene downstream from six copies of the SV40 late promoter region Tef-1 sites (SV40 nucleotides 200 to 219) (37); (ii) the promoter for the HCMV UL112-113 early gene; and (iii) the promoter for the HCMV late gene ICP36 (UL44). In order to generate dose-response data, we cotransfected U373MG cells with a constant amount of promoter-reporter plasmid (0.5 μ g) and increasing amounts (0, 0.1, 0.5, and 1.2 μ g) of the IEP86-expressing plasmids. The difference in input DNA concentrations was equalized with vector control plasmid pRSV3/Bg/II.

Figure 2 shows that all three forms of IEP86 activated each promoter relative to the promoters' basal activities (basal activity is set as 1 and is determined under conditions of transfection with 1.2 μ g of the vector control plasmid). Nevertheless, we noted significant differences in the levels of activation mediated by each form of IEP86. The fold activation of the Tef-TATA and ICP36 promoters by the Towne IEP86 was two to three times greater than that of the IEP86 produced from the original cDNA (OcDNA); further, the fold activation of

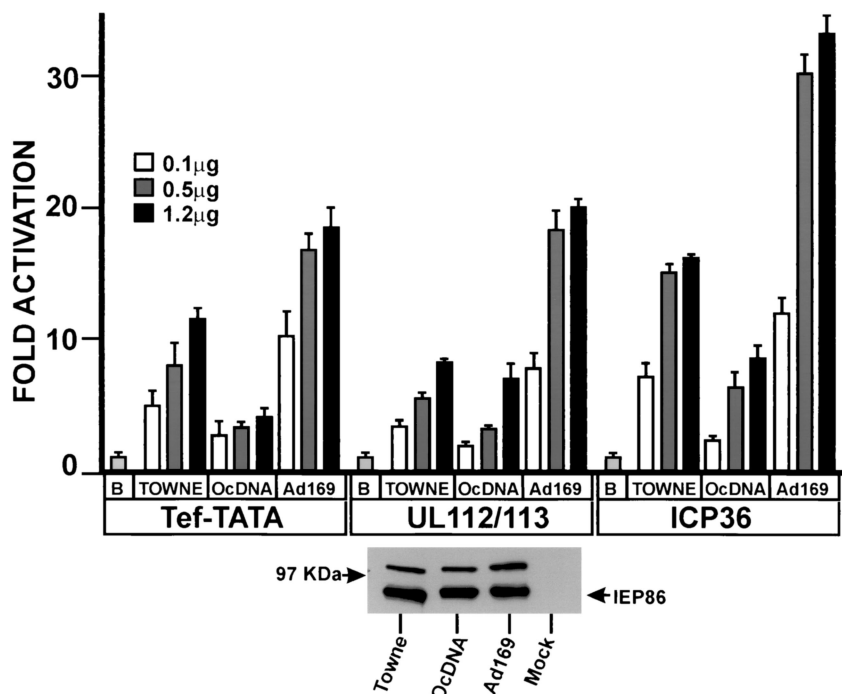


FIG. 2. Comparison of transcriptional activation mediated by Towne, AD169, and the OcDNA IEP86 on the simple Tef-TATA promoter and the promoters from the HCMV UL122-123 early gene and the late gene ICP36. U373MG cells were transfected with a constant amount (0.5 μ g) of luciferase reporter plasmid plus increasing amounts of IEP86-expressing plasmids. Basal promoter activity (B) is set to 1. The bottom panel shows a Western analysis of the three variants immunoprecipitated with MAb810 from U373MG cells transfected with each IEP86-expressing plasmid.

both of these promoters by the AD169 IEP86 was two to three times greater than that by the Towne IEP86. In contrast, with the UL112-113 promoter, the difference in activation between the Towne and OcDNA IEP86 was less significant, whereas the AD169 IEP86 again activated this promoter to levels two to three times greater than that of Towne IEP86. These data suggest significant strain differences in the activation of all three promoters and suggest, as in the case of UL112-113, that the amino acid variations and mutations affect some promoters more significantly than others.

The levels of expression of the different forms of IEP86 were repeatedly tested and found to be similar for all three. An example of such data is shown at the bottom of Fig. 2; total cell extracts were prepared from cultures transfected in parallel with those used in transcriptional activation studies, IEP86 was immunoprecipitated with monoclonal antibody MAb810, and IEP86 levels were determined by Western analysis as described in Materials and Methods.

In order to determine which of the amino acid variations and mutations mediate the differences in activity of the three forms of IEP86, we introduced each of the amino acid variations individually into the Towne background. Figure 1 shows each construction: T541A, A463T, K455E, M242I, and +S258 (insertion of an extra serine at 258). Figure 1 also shows various combinations of the amino acid substitutions which were constructed; these will be discussed below.

Effects of the different amino acid alterations on IEP86 expression and sumoylation. We first tested the levels of expression of the various IEP86 proteins containing the amino

acid variations. U373 cells were transfected for 48 h with 1.2 μ g of plasmids encoding each form of IEP86. Equal amounts of extract protein were immunoprecipitated with MAb810. The precipitated IEP86 was detected by Western analysis by probing with a polyclonal antibody made against the amino acids in exon 2 and 3 (anti-exon 2/3). Figure 3A shows an example of these data; although variations could be seen between individual experiments, the overall results from repeated experiments suggested that IEP86 variants containing the individual amino acid variations (T541A, A463T, K455E, and M242I) were expressed at nearly equivalent levels with the exception of A463T, which consistently produced IEP86 at higher levels (approximately twofold; see also Fig. 3B) than that of WT Towne IEP86. Combinations of amino acid variations between Towne and AD169 (K455E/T541A and K455E/T541A+S258) also demonstrated levels of IEP86 equivalent to that of WT IEP86.

Expression of the different forms of IEP86 was also tested by Western analysis of 45 μ g of total extracts (Fig. 3B). These data confirmed the relatively equivalent steady-state levels of IEP86 protein production between WT Towne and all the variants with the exception of A463T, which, as suggested in Fig. 3A, was increased.

Figure 3A and B also show that besides the major band of IEP86 (band A), there is a slower-migrating form (band B) which has been identified as being modified by small ubiquitin-like modifiers, namely, SUMO-1, -2, and -3 (1, 15). Sumoylation is a modification believed to increase the ability of IEP86 to activate different promoters (1, 15). Examination of the

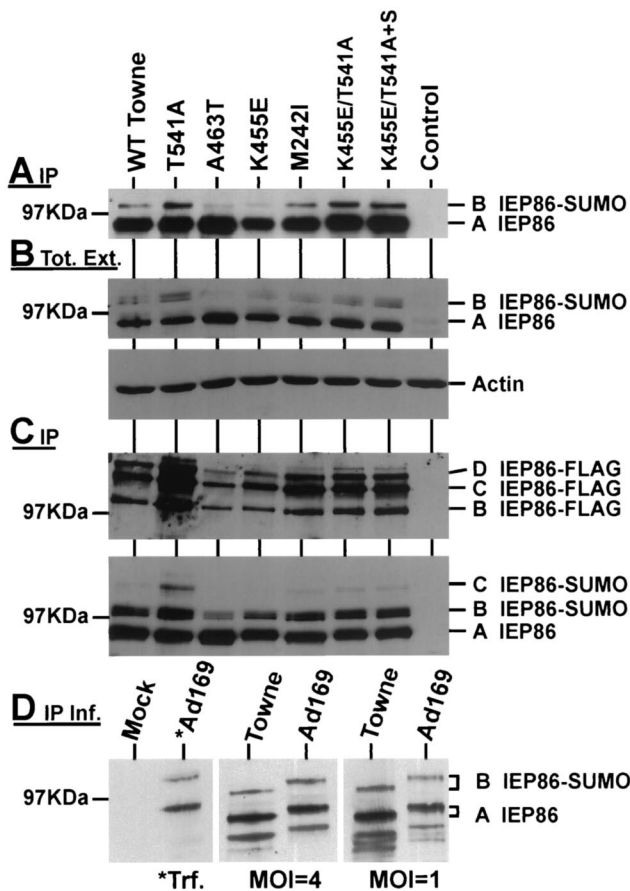


FIG. 3. Western analyses of IEP86 and sumoylated forms of IEP86. The Western analyses detect IEP86 in MAb810 immunoprecipitates (A, C, and D) or total extracts (B) from transfected or infected (D) U373MG cells. The U373MG cells were transfected with WT or variant IEP86-expressing plasmid (A), IEP86-expressing plasmid cotransfected with a Flag-SUMO-expressing plasmid (B and C), or infected with Towne and AD169 HCMV for 24 h at an MOI of 1 or 4 (D). The Western analyses were probed with anti-exon 2/3, an antibody which recognizes the common amino-terminal end of the MIEPs (A, B [top], and C [bottom]), anti-FLAG antibody (C [top]), anti-actin antibody (B [bottom]), or the exon 5 (IEP86)-specific antibody anti-pHM178 (D). Lane *Ad169 in panel D shows IEP86 produced in transfected (*Trf.) cells.

intensity of band B suggests that some of the amino acid variations altered the level of sumoylation. For example, T541A appears to increase the sumoylated IEP86 compared to WT Towne, whereas A463T and K455E appear to produce less of the sumoylated forms.

The extracts used to generate the data in Fig. 3B and C were derived from U373 cells which had been cotransfected with both the IEP86-expressing plasmids and with pHM976, a SUMO-1-expressing plasmid in which the SUMO moiety is fused to a FLAG epitope tag; thus, the sumoylated forms of IEP86 could be specifically detected with anti-FLAG antibody (15). The utilization of endogenous SUMO and FLAG-SUMO to modify IEP86 may account for band B being a doublet in Fig. 3B. In Fig. 3C, all forms of IEP86 were immunoprecipitated with MAb810 and analyzed by Western analysis by probing first with anti-FLAG antibody to detect the FLAG-sumoylated IEP86 (upper panel) and then with anti-exon 2/3 to

detect all forms of IEP86 (lower panel). The upper panel shows the multiply sumoylated forms of IEP86 that can be detected with the anti-FLAG antibody (bands B, C, and D). It is clear that the T541A alteration causes a significant increase in all forms of sumoylated IEP86. Interestingly, although K455E/T541A IEP86 and K455E/T541A+S258 IEP86 (see Fig. 1) each contain the T541A alteration, they do not show as great an increase in FLAG-sumoylated forms as the single T541A alteration. This implies that the addition of the K455E alteration may modulate sumoylation or that K455 is a sumoylation site; these implications are supported by examination of Fig. 3A and B, where the sumoylation of the K455E IEP86 appears to be moderately decreased compared to that of WT Towne. Finally, in Fig. 3A, B, and C, it is clear that the cDNA mutation, A463T, caused a significant decrease in sumoylated forms.

The bottom panel of Fig. 3C shows the same Western filter after stripping and reprobing with anti-exon 2/3. Here the predominant, nonsumoylated IEP86 (band A) can be seen as well as some of the sumoylated species (bands B and C). The slowest-migrating sumoylated forms are not well detected (bands C), or are not detected at all (band D), compared to their detection when probed with the anti-FLAG antibody (Fig. 3C). This has been seen by others who have suggested that these bands do not represent IEP86 (15). However, in Fig. 3C, the intensity of these bands in all samples varied proportionately with the known sumoylated form (band B), and thus we feel that bands C and D represent multiply sumoylated forms of IEP86. Due to the multiple sumoylation, the structure of these forms of IEP86 may be so altered that they are no longer easily detected by anti-exon 2/3 in Western analyses. Regardless, the data in Fig. 3C support the following conclusions regarding IEP86 sumoylation: (i) A463T significantly decreased the total sumoylation compared to WT Towne IEP86; (ii) T541A increased the total sumoylation compared to WT Towne IEP86; and (iii) the increased sumoylation seen with A451 is reduced by the K455E variation (we suspect that either E455 promotes conformational changes which cause decreased sumoylation at other sites or that K455 is itself a sumoylation site which is lost in the K455E variation).

The above data from the K455E variation suggest that total sumoylation of AD169 IEP86 may be quantitatively similar to the total sumoylation of Towne IEP86. Specifically, in AD169, A541 would increase sumoylation while E455 would decrease it, and in Towne, T541 would decrease sumoylation while K455 would increase it. In Fig. 3D, we have compared the sumoylation of the IEP86 from cells infected with Towne and AD169 viral stocks. U373MG cells were infected at MOIs of 1 and 4 and harvested 24 h after infection. The MIEPs were immunoprecipitated with mAb810, and IEP86 was specifically visualized by Western analysis by using the exon 5 (IEP86)-specific antibody anti-pHM178 (gift of T. Stamminger). Figure 3D shows that there is no quantitative difference in the levels of sumoylated IEP86 (band B) between the two strains at either multiplicity of infection. The second lane shows a sample derived from cells transfected with the AD169 IEP86 cDNA, marking the positions of IEP86 (band A) and the sumoylated form (band B). Figure 3D also shows that the AD169 forms of IEP86 migrate slower than those of Towne. This was also indicated in Fig. 2; however, the samples were not separated as

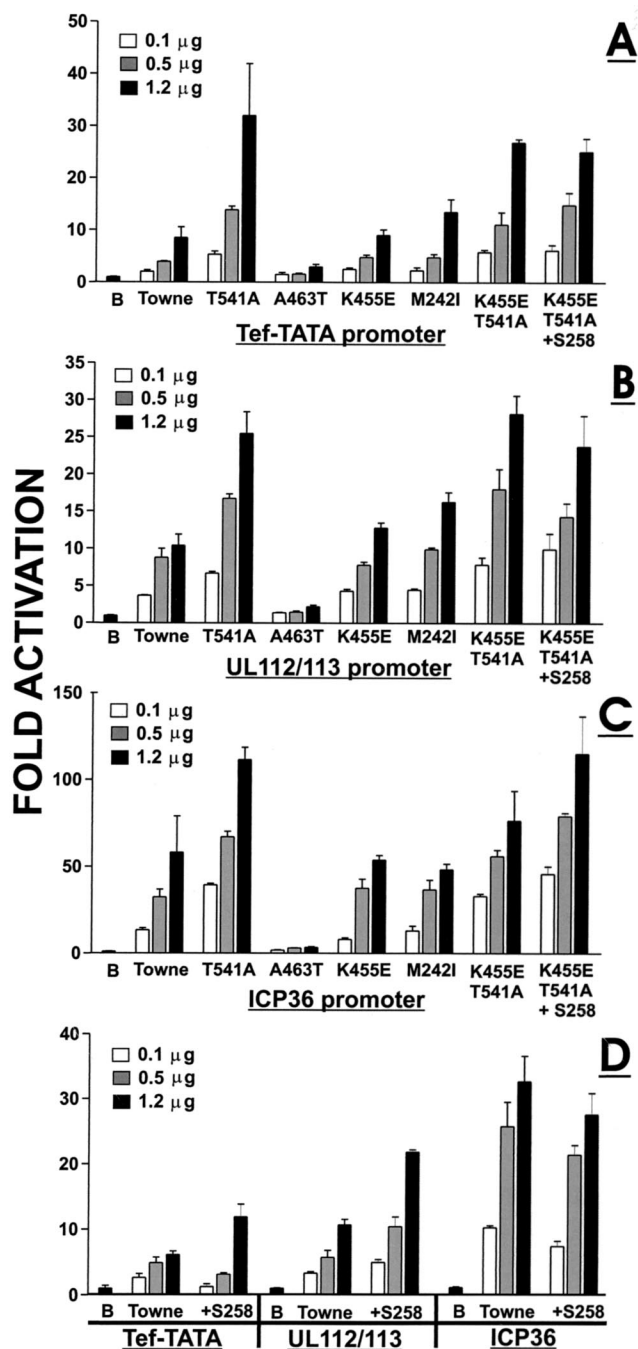


FIG. 4. Transcriptional activity of promoters by WT Towne IEP86 and IEP86 variants made in the Towne IEP86 background. Transfections and analyses were similar to those described in the legend for Fig. 2. The panels show activation of the simple Tef-TATA promoter (A), the UL112-113 promoter (B), and the ICP36 promoter (C). The effect of the +S258 IEP86 variant on each promoter is also shown (D). Basal promoter activity, shown in each panel by the column labeled with the letter B, is set at 1.

far as they were in Fig. 3D. The source of this migration difference is not known.

Transcriptional activation by IEP86 with the various amino acid alterations. We next assessed the ability of the various forms of IEP86 to transcriptionally activate by using the same

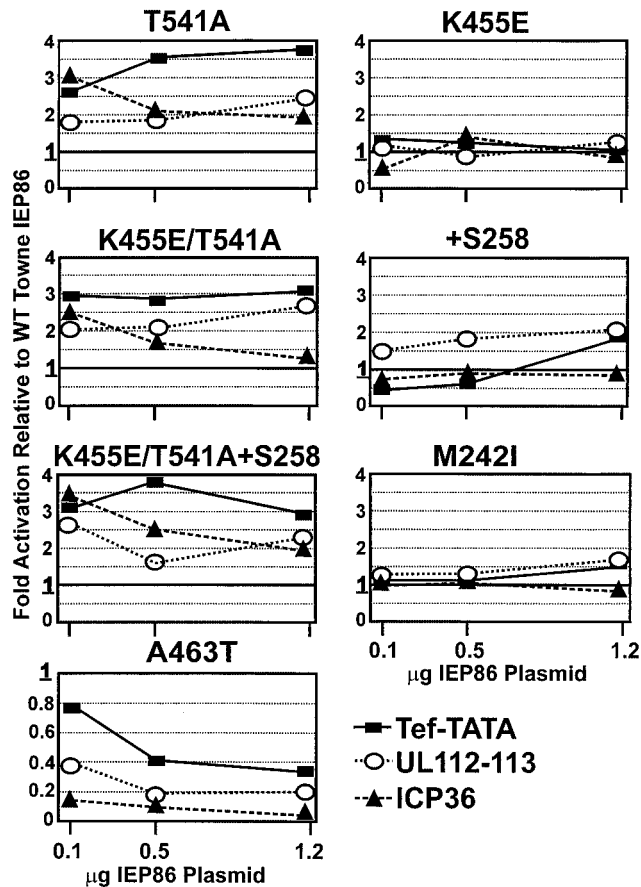


FIG. 5. Calculation of fold activation of the various promoters by the variant or mutant forms of IEP86 relative to activation by WT Towne IEP86 rather than to the basal promoter activity. Thus, the fold activation of the promoters by WT Towne IEP86 is equal to 1 at each input concentration of IEP86-expressing plasmid.

experimental approach described above. Figure 4A, B, and C show the data from experiments in which each form of IEP86 was tested on the three promoters Tef-TATA (Fig. 4A), UL112-113 (Fig. 4B), and ICP36 (Fig. 4C). Figure 4D shows the activation of the three promoters with the +S258 IEP86 (Fig. 1); these data were gathered in separate experiments. In Fig. 4A to D, fold activation is calculated relative to the basal promoter activity (B), which is set at 1. In Fig. 5, the same data were recalculated to show fold activation of each IEP86 variant relative to the activity of WT IEP86 at each input plasmid concentration; in this case, the activity of WT IEP86 is always equal to 1. This comparison, presented as line graphs, shows the effects of each individual form of IEP86 comparatively on each promoter.

T541A represents one of the sequence variations between the AD169 and Towne strains; T541A IEP86 was more highly sumoylated (see above). The T541A IEP86 transcriptionally activated each of the promoters two to three times better than WT Towne IEP86 (Fig. 4A, B, and C). The increased transcriptional activation was most striking with the Tef-TATA promoter (Fig. 4A). This is also seen in Fig. 5, where T541A activation of the Tef-TATA promoter increased relative to that of WT Towne IEP86, with increased amounts of the

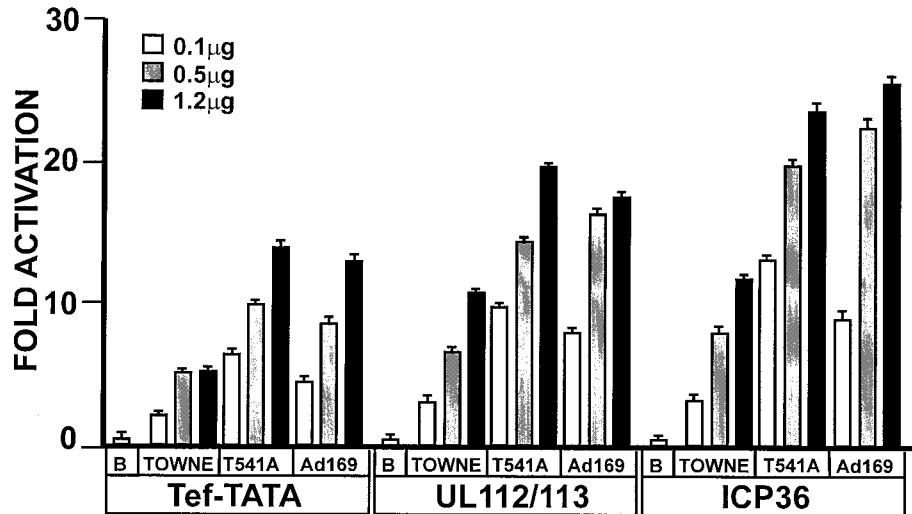


FIG. 6. Transcriptional activity of the T541A IEP86 variant relative to those of WT Towne and WT AD169. The T541A variation accounts for the greater activity of AD169 IEP86 relative to that of Towne IEP86. Transfections and analyses were similar to those described in the legend for Fig. 2. Basal promoter activity, shown in the columns labeled with the letter B, is set to 1.

IEP86-expressing plasmid; however, the opposite was seen with the ICP36 promoter, where fold activation relative to WT decreased with increased expression plasmid input. These data suggest that the intracellular level of T541A IEP86 may differentially affect specific promoter activation.

K455E represents another sequence variation between the AD169 and Towne. K455E IEP86 showed modestly decreased sumoylation compared to WT Towne IEP86. The data in Fig. 4A, B, and C show that transcriptional activation of each promoter by K455E IEP86 was very similar to that of WT Towne IEP86. This is also shown in Fig. 5, where fold activation of each promoter by K455E remains close to 1, suggesting that it functions like WT Towne IEP86.

+S258 represents the variant with the extra serine found in AD169 within the stretch of serines between 258 and 264. Figure 4D shows that +S258 IEP86 produced a modest increase in transcription activation (twofold or less) on the Tef-TATA and UL112-113 promoters. This effect on the Tef-TATA promoter was seen only at the highest input (1.2 µg) of the +S258 IEP86 expression plasmid. Interestingly, the +S258 IEP86 showed slightly decreased activation of the ICP36 promoter. These findings, suggesting modest preference for the UL112-113 promoter, can be better seen in Fig. 5.

K455E/T541A and K455E/T541A+S258 are forms of IEP86 made to test the effect of combining the sequence differences between the Towne and AD169 strains. The comparison of the activation of the different promoters by K455E/T541A, K455E/T541A+S258, and T541A is best seen in Fig. 5, where activation is shown relative to that of WT Towne IEP86. First, it can be seen that the effects of all three forms of IEP86 are generally similar on each promoter. Specifically, (i) the Tef-TATA promoter is always activated to the greatest extent, and activation often increases with increased expression plasmid input; (ii) activation of the ICP36 promoter decreases with increased expression plasmid input; and (iii) activation of the UL112-113 promoter is very similar at all input expression plasmid amounts. Overall, these data suggest that among the amino

acid differences between Towne and AD169, the T541A variation alone has the most significant effect in these studies. However, more subtle effects can be seen with the other variations; examination of the data in Fig. 4C shows that K455E/T541A showed decreased activation of the ICP36 promoter compared to T541A IEP86. Interestingly, the addition of an S at 258 (K455E/T541A+S258) restored much of this lost activation.

AD169 IEP86 was obtained by the addition of the R68Q alteration to K455E/T541A+S258 (Fig. 1). Numerous experiments have shown little difference between the activities of K455E/T541A+S258 IEP86 and the fully AD169 IEP86. Thus, we conclude that the R68Q variation does not significantly affect transcriptional activation of these promoters. The cumulative data suggest that the T541A alteration is the primary

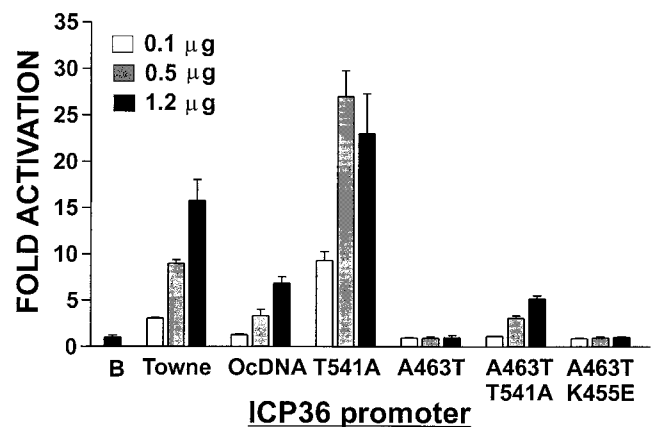


FIG. 7. Transcriptional activity of IEP86 variants relative to that of WT Towne. The T541A variation can overcome some of the negative effects of the A463T mutation, allowing the OcDNA IEP86 to have some transcriptional activation function. Transfection conditions and analyses were similar to those described in the legend for Fig. 2.

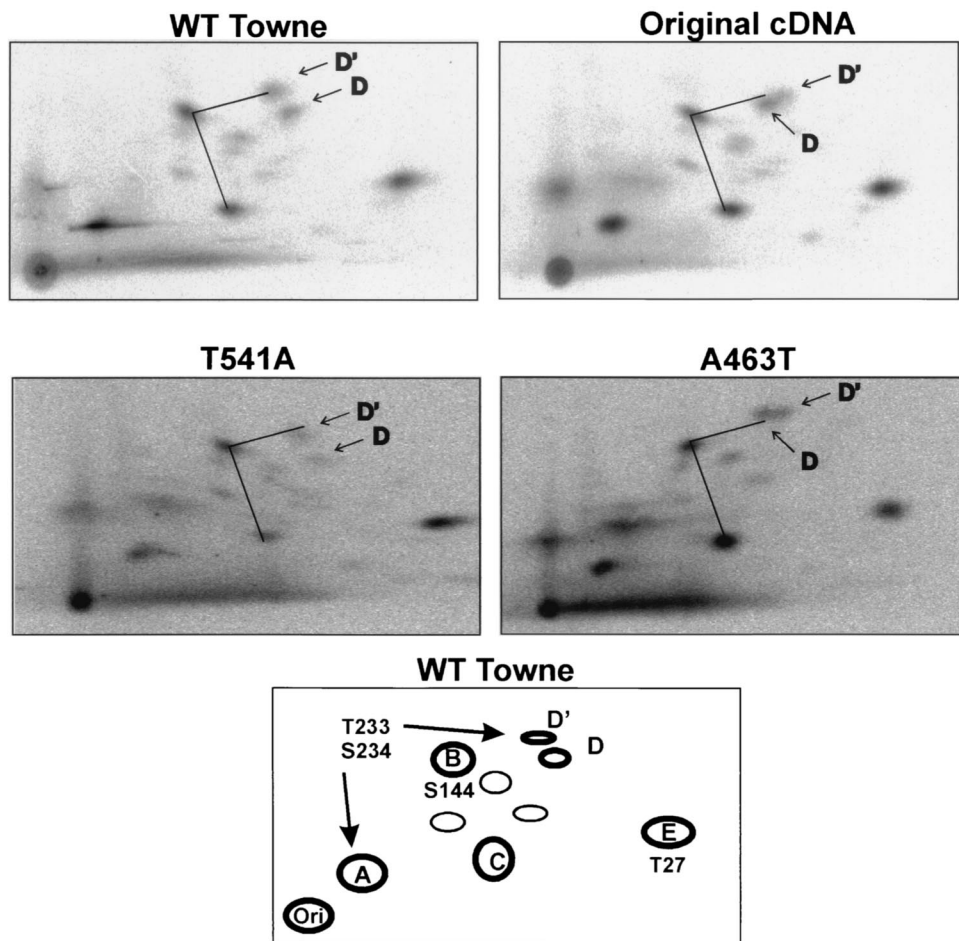


FIG. 8. Tryptic phosphopeptide maps of WT Towne, OcDNA, T541A, and A463T IEP86. The bottom panel shows a schematic phosphopeptide map of WT Towne with the prominent phosphopeptides labeled A, B, C, D/D', and E (12). Mutations affecting some of the peptides are shown. For easier comparison of the data, identical lines are drawn connecting phosphopeptides C, B, and D' in each panel.

reason that AD169 IEP86 is a better transcriptional activator than Towne IEP86. In Fig. 6, we show this by direct comparison of the transcriptional activation of WT Towne, T541A, and WT AD169 IEP86 on each promoter. The data clearly show that the T541A IEP86 activates each promoter to approximately the same extent as WT AD169 IEP86.

M242I and A463T represent the two amino acid differences in the original cDNA which are found in neither the Towne or AD169 sequences (Fig. 1). Figure 4A, B, and C and Fig. 5 show that M242I is a silent mutation with respect to transcriptional activation of the test promoters. It was also silent with respect to sumoylation (see above). However, the data in Fig. 4A, B, and C show that the A463T alteration causes a dramatic decrease in transcriptional activation of all the promoters. Figure 5 shows that relative to WT Towne IEP86, activation is decreased at every input amount of expression plasmid and that the effect is greatest at the higher input amounts. Recall that A463T IEP86 is increased about twofold in expression compared to other forms of IEP86 and is the least sumoylated of any of the various IEP86 variants (Fig. 3A, B, and C).

Interestingly, the OcDNA IEP86 contains the deleterious A463T mutation (Fig. 1) yet retains an ability to transcription-

ally activate the test promoters (Fig. 2). In repeated experiments, we have noted that it is the presence of the T541A mutation which partially compensates for the deleterious effects of the A463T mutation. Figure 7 shows an example of these data comparing the activation of the ICP36 promoter by WT Towne, OcDNA, T541A, A463T, A463T/T541A, and A463T/K455E IEP86. It is clear that the addition of T541A to A463T restores some transcriptional activity such that A463T/T541A IEP86 has essentially the same level of activity as OcDNA IEP86. In a control experiment, the combination of A463T and K455E did not affect the inhibition of activation seen with A463T alone. Similar results have been seen with the Tef-TATA and UL112-113 promoters (not shown). Thus, it appears that the OcDNA IEP86 retains transcriptional activity primarily because of the T541A alteration.

Phosphorylation of the various forms of IEP86. Previous experiments have examined the tryptic phosphopeptide map of IEP86 and have documented prominent peptides which were designated A, B, C, D, D', and E (see the schematic in Fig. 8) (12). Previous data showed that alanine mutations of specific mitogen-activated protein kinase sites affected phosphorylation of some of the peptides: S144A eliminated phosphoryla-

tion of peptide B, T27A eliminated phosphorylation of peptide E, and T233A/S234A had multiple effects eliminating phosphorylation of peptides D and D' and reducing phosphorylation in peptide A (12).

In the preset studies, the two amino acid substitutions which caused the greatest effects, T541A and A463T, could each affect phosphorylation of IEP86; threonine at position 541 does not form a predictable phosphorylation motif, and threonine at position 463 is included in a CKII consensus site. Tryptic phosphopeptide maps of WT Towne, the OcDNA, T541A, and A463T IEP86 are shown in Fig. 8. The tryptic maps of WT Towne IEP86 and the OcDNA IEP86 are very similar except for the migration of phosphopeptide D; the identical angled line between phosphopeptides C, B, and D' is drawn in each map for easy comparison. T541A is also very similar to WT Towne, suggesting that T541 is not a major phosphorylation site. The A463T mutation caused an altered migration of phosphopeptide D, similar to that seen with the OcDNA IEP86, suggesting that A463T is responsible, at least in part, for this difference. It is interesting that the mutation T233A/S234A eliminates phosphorylation of peptide D (12) while A463T, a distant amino acid, affects its migration (see Discussion).

DISCUSSION

Effects of amino acid variations and mutations on promoter activation. R68Q, +S258, K455E, and T541A represent sequence variations between the AD169 and Towne strains. We have not directly tested the R68Q variation but note that there is no significant difference in transcriptional activation of the test promoters between WT AD169 IEP86 (containing Q68) and K455E/T541A+S258 IEP86, which is identical to AD169 IEP86 except that it contains R68 (Fig. 1). The addition of a serine at 258, +S258, showed modestly increased transcriptional activation, which was preferential for the UL112-113 promoter, while the other two promoters were generally less activated relative to activation by WT Towne IEP86. K455E IEP86 functioned much like WT Towne IEP86 in transcriptional activation. However, the T541A alteration significantly increased transcriptional activation on all promoters tested. T541A appears to be the major alteration making WT AD169 IEP86 a better transcriptional activator than WT Towne IEP86. Promoter specificity was suggested in Fig. 5, which shows that increased input of the T541A IEP86-expressing plasmid resulted in greater activation of the Tef-TATA promoter but decreased activation with the ICP36 promoter. This suggests that the intracellular level of IEP86 may differentially affect specific promoter activation.

M242I and A463T represent the two amino acid mutations in the OcDNA (Fig. 1). M242I was found to be a silent mutation in our analyses. A463T IEP86 was dramatically deleterious to transcriptional activation of all promoters; however, it was generally more detrimental to the viral promoters than to the Tef-TATA promoter. This deleterious effect on transcriptional activation can be partially offset by T541A, thus accounting for the ability of the OcDNA IEP86 to retain any ability to transcriptionally activate.

In sum, the individual effects of the several amino acid variations show that (i) AD169 IEP86 is, in general, a better

transcriptional activator than Towne IEP86, (ii) some amino acid variations affect promoters preferentially, suggesting that IEP86 may have separate domains, or conformations, which differentially affect specific promoters, and (iii) the intracellular concentration of IEP86 affects promoters differentially, which would be significant as the intracellular concentration of IEP86 changes during a lytic infection.

Effects of amino acid variations and mutations on sumoylation. Sumoylation occurs on lysines, and there are two major alternative sumoylation sites known in IEP86, lysines 175 and 180 (15). Sumoylation of IEP86 has been associated with increased transcriptional activation. Specifically, Hofmann et al. (15) reported that an IEP86 sumoylation-defective mutant (mutated at both lysines 175 and 180) showed significantly reduced transcriptional activation of two viral promoters (UL84 and UL112-113). In addition, Ahn et al. (1) reported that cotransfection of plasmids encoding WT IEP86 with the SUMO-conjugating enzyme (Ubc9) leads to increased transcriptional activation of the CycE promoter, apparently due to increased ability to sumoylate IEP86. Correspondingly, this increased activation was not seen when the sumoylation-defective mutant was used.

T541A IEP86 shows increased sumoylation, which may account, at least in part, for its increased transcriptional activation. The effect of T541A on sumoylation may be to alter the structure of IEP86 such that the major sites are better utilized or so that new sites are exposed. This increased sumoylation is not seen when the K455E alteration is combined with T541A, leading to the speculation that K455 may be an alternate sumoylation site lost with the change to E455 or that E455 alters the structure of IEP86 such that sumoylation is modulated at other sites in the protein. The correlation between the amount of sumoylation and transcription is not as apparent when comparing T541A/K455E IEP86 with T541A IEP86. Despite less sumoylation of the T541A/K455E IEP86, the effects on promoter activation showed only subtle differences, primarily on the ICP36 promoter, where K455E/T541A showed modestly decreased activation compared to T541A IEP86 (Fig. 4C). However, we stress that our measurements of sumoylation are quantitative, not qualitative. Thus, while the increased total sumoylation seen with the A541 variation is down-modulated by E455, the sumoylation of the sites which most critically affect transcription may be unaltered. In this regard, the effects of the K455E variation on sumoylation lead to the prediction that total sumoylation may be quantitatively similar between Towne and AD169 IEP86. We have shown that this is the case in viral infections of U373MG cells. However, this quantitative similarity in sumoylation between the two strains cannot be used as an argument against a role of sumoylation in the different levels of transcriptional activation seen between Towne and AD169 IEP86. As argued above, although the quantity of sumoylated forms is similar, the sites of sumoylation may be very different. Further study is needed to settle this matter.

A correlation between sumoylation and transcriptional activation was seen with A463T IEP86. Here we noted quantitatively decreased sumoylation and dramatically decreased transcriptional activation. The question from these data is whether the decrease in sumoylation of A463T IEP86 can fully account for the dramatic reduction of transcriptional activation of all

promoters tested. As discussed below, we feel that the A463T mutation significantly alters the functional structure of the protein, affecting not only sumoylation but also phosphorylation and the structural integrity of the carboxy-terminal region of the protein.

Structural analysis. The phosphopeptide mapping data showed that the A463T mutation resulted in a change in the mobility of phosphopeptide D. Since amino acid 463 is not within peptide D (it contains amino acids surrounding T233/S234), this observation suggests that IEP86 may have a dynamic conformation in which the A463T mutation alters the conformation such that posttranslational modification of amino acids within phosphopeptide D is altered. Similarly, the effects of A463T, T541A, and K455E on sumoylation, primarily at lysines 175 and 180, suggest a dynamic protein conformation.

In order to predict how A463T and T541A might be involved in conformational alterations in IEP86, we performed secondary and tertiary structural modeling. The fact that IEP86 does not have enough similarity to proteins of known structure eliminated the use of homology modeling. Thus we chose to model by using three programs, SAM-T99, PSIPred, and PROFsec (see Materials and Methods), which have been shown to perform well when there are several homologous variations of the protein of interest. In this regard, the C-terminal portion of IEP86 between amino acids 344 and 579, which contains K455E, A463T, and T541A, is highly conserved among IEP86-like proteins encoded by beta-herpesviruses (Fig. 9). In addition, the C-terminal region contains independent functions, such as the dimerization domain (amino acids 390 to 542) and much of the DNA binding domain (amino acids 330 to 579) (7).

The PROFsec, PSIPred, and SAM-T99 analyses each predicted that the C-terminal region contained mostly alpha helices. PROFsec prediction for globularity showed this domain as globular. Due to the similarity of the analyses, only the PROFsec results are shown in Fig. 9 and 10.

Structural predictions for A463T and T541A. Figure 10 highlights the regions immediately around A463 and T541. A463 is predicted to be buried, forming part of an amphipathic helix. This would agree with the observation that there is always a hydrophobic amino acid (either alanine or valine) in that position in IEP86-like proteins from all the beta-herpesviruses (Fig. 9). The A463T mutation puts a polar amino acid in this position, resulting in lowering of the reliability of the helical prediction and changing the predictions for solvent exposure. These alterations would be expected to have significant effects on interactions with nearby amino acids. This will be examined by three-dimensional modeling below.

T541 is predicted to be in a weakly helical region, and the alteration to A541 significantly strengthens the helical prediction in this region and alters solvent accessibility, which may account, in part, for the positive effect of the T541A alteration.

Three-dimensional modeling. In addition to the PROFsec modeling, we modeled the three-dimensional structure of amino acids 344 to 579 by using THREADER and other programs as described in Materials and Methods. THREADER is based on the observation that proteins with no sequence similarity can show similarities in native folding. The THREADER approach is not based on sequence but uses pairwise potentials with comparison to a library of protein

folded derived from established protein structures. The THREADER three-dimensional model predicts a globular structure consisting of eight alpha helices (Fig. 11A and B). Comparing the location of the helices in this model with the locations of helices in the two-dimensional model (Fig. 9) shows good correlation. For comparison, the alpha helices predicted on the three-dimensional model are indicated in Fig. 9 by the thick lines labeled helix 1 through 8 (H1 to H8). Five helices overlap in both models, and two partially overlap. While three-dimensional modeling is highly speculative (hence its presentation in Discussion), the similarity of the predictions from two- and three-dimensional modeling, two very different structure analysis approaches, strengthens the reliability.

In Fig. 11A to F, we show several representations of the proposed model. Figure 11A shows the ribbon structure in spectral colors going from the N terminal (blue) to the C terminal (red). Amino acids significant for this discussion (T541, A463, C464, and C509) are highlighted as ball-and-stick models. Figure 11B, C, and D show different views of the space-filling model of the structure, in which hydrophobic amino acids are shown in orange, A463 is shown in red, C464 and C509 are shown in black, and T541 is shown in green. A463 is part of helix 5, and, as indicated in Fig. 11A, B, and C, it is partially buried, as predicted above. The structure of the amino acids located within 4 Å of A463 is shown in Fig. 11E. These include amino acids 459 to 467 in helix 5, plus threonine 440 and leucine 443 from helix 4; hydrophobic amino acids are shown in orange. L443 is modeled to be very close to A463 due to their shared hydrophobicity (carbon beta of A463 is 3.4 Å from the carbon delta 1 of L443). The potential structural significance of L443 is indicated by studies (38) showing that the mutation of amino acids 442 to 444 (FLM to ASA) eliminated dimerization and DNA binding. Introducing a polar amino acid, e.g., threonine, at position 463 would decrease the affinity and proximity of residues 463 and 443, since they would not cluster together. This change would result in an alteration in the interaction between helices 4 and 5. The deleterious effects of the A463T mutation, seen experimentally, suggest that maintenance of the hydrophobic clustering of the amino acids shown in Fig. 11E may maintain a core structure within the carboxy-terminal half of IEP86.

Further indication of the importance of A463 in the maintenance of IEP86 structure comes from the observation that its adjacent amino acid, 464, is a cysteine. In Fig. 11A, B, and C, C464 (black) is shown to be buried. Interestingly, Fig. 11A shows that C464 faces C509 located in helix 6. Figure 11F expands this region, detailing the positions of A463, C464, and C509, and suggests the potential for a disulfide bond. Mutation of C509 (C510 on the AD169 strain) to glycine results in a temperature-sensitive mutation (13), suggesting that a disulfide bond involving C509 may be structurally important. Clearly, the disruption in structure caused by the A463T mutation, described above, may alter the position of C464 such that it would no longer be in position to form a disulfide bond with C509. If this was the case, then the A463T mutation would alter the way helix 5 interacts with both helix 4 and 6.

Modeling does not as clearly indicate why the T541A alteration increases transcriptional activation and sumoylation of IEP86. T541A is an exposed residue in helix 7 (Fig. 11A, B, and D), and changing it to an alanine is not predicted to have a

| | | | | | |
|-------------------|--|---|--|----------------|-----|
| IEP86 AA Number | 344 | <u>H1</u> | <u>H2</u> | <u>H3</u> | 412 |
| HCMV Towne IEP86 | K | NTPFCTPNVQTRRGRVKIDVSRMFRMNTNRSLEYKLNLPFTI | PSMHQVLDEAIKACKTMQVNNKGIQII | | |
| Simian CMV IE2 | KGGVTF | TNPVKVETKRGRVQDDVSRMFRMNTNRSLEYKLNLPFQPTNVHQILSEAVTVCKTMQVSNRGMILI | | | |
| M.Mulatta CMV IE2 | KAGTVFKNPTVETKRGRVRADEVSRMFRATRSLEYKLNLPFLTTNVHQLMAEAVNACKTMQVNHARGIMVT | | | | |
| Tupaia HV T22 | KT.QPFDK | PQGESKRGRVKSGLDLRLFRTRTQRSLEYKTLPPKAPPLHEVWEQTVRACRELOVLNKAILLS | | | |
| Rat CMV IE2 | ..GPGFV | APNAHKRGKRNREGNNDGRGKPTRALEYKQMPYKQQTQVQFLYGNAIARTCRESTVHDKIIMVM | | | |
| Murine CMV IE3 | RTAGGYV | APNAHKCR.....EDKSRKYPARALEYKLNLPFRPQSPQYLLGKAIQFCKEETVHDKFIMLF | | | |
| HHV7 IEA | QEQK..QSAKFKVQRGRVPTKDFEKLFRKTIRAFEYKQIPKPFPEAKLKEVSVALCCNGSSSKGALIIY | | | | |
| HHV6 MIE-2 | KERK..VSEKFKIHRGRVATKDFQKLFRTMTRAFEYKQIPKPCNEKNLKEAVYDICCNGLSNNAAIMY | | | | |
| PROF Sec. Struct. | L LLL LL LL L | HHHHH HHHHH LLLLLLHHHHHHHHHHHHH | L EEEEE | | |
| Reliability SS | 9 | 87532564221255335430114310010112201477667876888999999885212345347898 | | | |
| Solv. Access | e | eee e ee eeeee eeee eebbeb bbeb e e ee eebbeebbee ebeb bbbb | | | |
| Reliability Acc. | 7 | 22201112010110103032012012230231211221020120235033804192222020139896 | | | |
| K455 A463 | | | | | |
| IEP86 AANumber | 413 | <u>H4</u> | <u>H5</u> | 481 | |
| HCMV Towne IEP86 | YTRNHEVK | EVDVAVRCLGTMCNLALSTPFLMEHTMPVTHPKVAQRTADACNEGVKAAWS.LKELHTHQ | | | |
| Simian CMV IE2 | YTRTHEVKEAVD | TARIRLGGICNLAI | STPFLMEHTMPVHNPEITRKTAEACQOQLRAVWD.LKTVQPHN | | |
| M.Mulatta CMV IE2 | YTRTHEVKEAVD | VARVKLGKIPNLSIS | TPFLVEHTMPNVYPAEVVRKTAECASGKQAWD.LKAVQPHD | | |
| Tupaia HV T22 | YTRTYD | VRSAVEDVRAKLGTTVNLIS | SAPWMEHTLPRVHTPQTVSRACRDGVKTAWE.LQEHHTHQ | | |
| Rat CMV IE2 | YTRGQDI | RQAEIEKLRSLQGIITNLS | SAPFTEHTKQIHTPNTVNMTSQALAAAGLQASWN.LDEDNKHN | | |
| Murine CMV IE3 | YTRSQD | VRKAVIDETRARMGMRPNLS | ISCPFMEHTKPINHSRETIDRTSACTAGTQAVWD.MEERRGQK | | |
| HHV7 IEA | FTRSKT | VAEDIKAMQKELMVRPNTIMSEPFKMNHAPPRYYDKDSIKNFIELQKKGPHLEWENT.SNSAHT | | | |
| HHV6 MIE-2 | FTRSKKVAQIIKIMQKELMIRPNITVSEAFKMNHAPPKYDKDEIKRFIQLQKQGFQELWDKFENNTTHD | | | | |
| PROF Sec. Struct. | EE LHHHHHHHHHHHHH LL EEE | LLLLLHHHHHHHHHHHHL | HHHEE HHH LLLL | | |
| Reliability SS | 5045457989899998724654213304111234777678846788889988605111100 | 11025566 | | | |
| Solv. Access | bb | eb eebbeebbeb e bbbbbbbbbb be e e e ee be bbebbebbe bbe bee ee e | | | |
| Reliability Acc. | 6402125231633701330001133430120021010011125033352243330031202 | 14211211 | | | |
| T541 | | | | | |
| IEP86 AA Number | 482 | <u>H6</u> | <u>H7</u> | 550 | |
| HCMV Towne IEP86 | LCPRSSDYRNMI | IHAATFVDLLGALNLCLEPLMQKFPKQVMVRI | FST.NQGGFMLPIYETATKAYAVGQFE | | |
| Simian CMV IE2 | LCPRSSDYRMI | IINSATFVDFLAAIQVLIPLVQYQPKQVAIRIFSAENTS | SSFMLPIYDNASRMYAVGQFE | | |
| M.Mulatta CMV IE2 | MCPRSSDYRNMI | VHAATFVDLGLSLOVLVPLVQRFPKQVAIRIFTNENSNS | FMLPIYDQAAKMYAVGQFE | | |
| Tupaia HV T22 | YCPRASDVRT | IMIQAATPCEFLTASRLCLELMQKYPKQVAVRTCTIDGGHN.LLPIYEGPSSDYSNQQFD | | | |
| Rat CMV IE2 | NAPRMSDYRMI | IQAATPPDFLGALKLCIQFAQTFPKNACIRLCNIVGGLQ.PLPPIYEVVYATDQYN | | | |
| Murine CMV IE3 | CVPRTSDYRSMIIQAANPPDFLGA | VKTCLHLSQVFPKQVCMRLCSITGGLN.PLPPIYEETVSSYVNAQFE | | | |
| HHV7 IEA | LYTRHSD | LKTI IYAATPIDFFMAKICNNYAKRPKEIVLRLSSITDGN.PISINPVSRDFL.SKYT | | | |
| HHV6 MIE-2 | LFTRHSDVKTMIYAATPIDFVGA | VKTCNKYAKDNPKIEIVLRVCSIIDGN.PISINPVISKEFK.SKFS | | | |
| PROF Sec. Struct. | LLLL | EEEE L HHHHHHHHHHHHHHLL EEEEEEE | LLL | HHHHHHH | |
| Reliability SS | 666543100022221462458888888887631664258887622 | 77642100113012453101101 | | | |
| Solv. Access | b | b b bbbbbbbbbb bb bbebb ebee ee bbbbbbbb e eebbbb eebbebbe ebe | | | |
| Reliability Acc. | 0001001102779164010163185329030435112256918530 | 10003131500011210010202 | | | |
| IEP86 AA Number | 551 | <u>H8</u> | 579 | | |
| HCMV Towne IEP86 | QPTETPPEDLD | TLSLAEAAIQDLRNKSQ~~~~~ | | | |
| Simian CMV IE2 | EPKD...ELENLS | MATEAATDM.QEDSQ~~~~~ | | | |
| M.Mulatta CMV IE2 | ETKD...EDLANLS | MATEKAIEDMNQESSQ~~~~~ | | | |
| Tupaia HV T22 | VQTN | SCNR.LDEINEAVAQAIRDLESGDREQQQQQ | | | |
| Rat CMV IE2 | FSPITNKDS | NGMSTILDQD.SDSE~~~~~ | | | |
| Murine CMV IE3 | ADDISHHE | DESGL...EYE.SDCE~~~~~ | | | |
| HHV7 IEA | ALCKF | ~~~~~ | | | |
| HHV6 MIE-2 | TLSKC | ~~~~~ | | | |
| PROF Sec. Struct. | LLLLL | HHHHHHHHHHHHHHH | | | |
| Reliability SS | 47778721033356889888877521434 | | | | |
| Solv. Access | e | eeeeeeeeebeebebebebebebe | | | |
| Reliability Acc. | 00230123102241258135633022324 | | | | |

FIG. 9. Secondary structure and solvent accessibility prediction of the C-terminal end of IEP86. The structural predictions were made by using the program PROFsec and MSA, as described in Materials and Methods. Amino acids in the various β -herpesvirus IEP86-like proteins which are identical to the HCMV Towne IEP86 sequence are shown in bold. Shown at the bottom of each MSA block are the predictions for secondary structure (PROF Sec. Struct.) (H, alpha helix; E, beta strand; L, loop) and solvent accessibility (Solv. Access) (e, exposed; b, buried) and the reliability of these predictions (Reliability SS and Reliability Acc., respectively). The reliability of the predictions varies between 0 (low) and 9 (high). Secondary structure predictions made with an expected average accuracy higher than 82% are in bold. The subset of solvent accessibility predictions with reliability higher than 4 is shown in bold. For a comparison, the alpha helices predicted for the three-dimensional model (Fig. 11) are shown as thick lines at the top of each block (H1 to H8).

| | | |
|---------------------|----------------------------------|---|
| WT IEP86 | A463 | T541 |
| IEP86 AA Number | 454 ▼ 477 | 528 ▼ 550 |
| HCMV IEP86 IE2 | PKVAQRTADACNEGVKAAWS.LKEL | NQGGFMLPIYETA TKAYAVGQFE |
| PROF Sec. Struct. | HHHHHHHHHHHH HHHEE HHH | HHHHHHH |
| Reliability SS | <u>46788889988605111100</u> 1102 | 77642100113012453101101 |
| 82% Subset SS | .HHHHHHHHHH.L..... | LLL..... <u>H</u> |
| Solv. Access | ee be bbebbbee bbe bee | e eebbbb ebbeebe ebe |
| Reliability Acc. | 25033352243330031202 1421 | 10003131500011210010202 |
| Subset Acc. | <u>.e...b..b...</u> e.. |b..... |
| MUTANT IEP86 | T463 | A541 |
| PROF Sec. Struct. | HHHHHHHHHHHH HHHHH HHH | HHHHHHHHH |
| Reliability SS | <u>35677678777513133200</u> 1102 | 7765323001256775311112 |
| 82% Subset SS | .HHHHHHHHHH..... | LLL..... <u>HHHHH</u> |
| Solv. Access | ee beebbebbeebbe bbe bee | eeeebbbbb ebbe e ebe |
| Reliability Acc. | 25033231323342030202 2321 | 11011020422235420011443 |
| Subset Acc. | <u>.e.....e.....</u> |b..... <u>be</u> <u>eb</u> . |

FIG. 10. Comparisons of secondary structure and solvent accessibility predictions for the changes A463T and T541A. Structure predictions based on PROFsec and MSA were prepared as described in the legend for Fig. 9. For the T463 prediction, a T was placed at position 463 in all the sequences in the MSA; likewise, an A, instead of T or S, was inserted at position 541 for the A541 prediction. Prediction abbreviations are described in the legend for Fig. 9, with the exception of Subset Acc. (subset of solvent accessibility predictions with reliability higher than 4) and 82% Subset SS (secondary structure predictions made with an expected average accuracy higher than 82%). Changes in the predictions resulting from the mutations are underlined.

dramatic structural effect according to this model. However, we note that A541 would increase the hydrophobic surface (Fig. 11D), possibly making the region more efficient in protein-protein interactions, including dimerization, and/or intramolecular interactions with the N-terminal half of the pro-

tein, which could account for its effects on sumoylation. Such effects on intra- and intermolecular interactions by the T541A variation may account for how it partially compensates for the A463T mutation through the resurrection of some of the lost structure caused by T463.

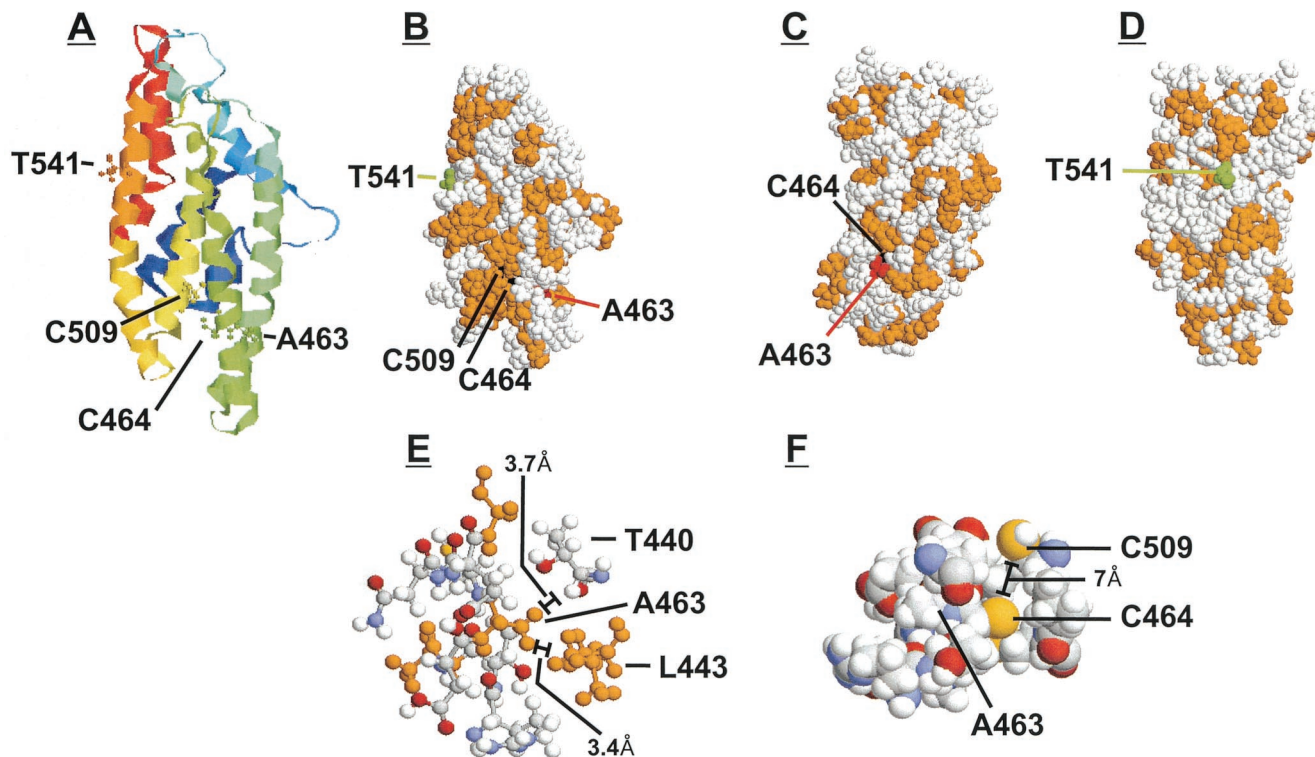


FIG. 11. THREADER three-dimensional model of the IEP86 C-terminal fragment. Three-dimensional structures were predicted by using THREADER version 2.5 (17). (A) Shown is the ribbon structure colored from blue to red in the direction from the N terminus to the C terminus. Amino acids 463, 464, 509, and 541 are highlighted as ball-and-stick models. (B, C, and D) Shown are space-filling models where hydrophobic residues are colored in orange, T541 is shown in green, C464 and C509 are shown in black, and A463 is shown in red. (E) Shown are amino acids located 4 Å or less from A463. Hydrophobic residues are orange. (F) Shown are amino acids located 5 Å or less from C464.

The finding that amino acid alterations and mutations known to have significant functional effects also have significant predicted effects on the modeled protein lends credence to the structural predictions and three-dimensional modeling. We feel that IEP86 may have a dynamic tertiary structure in which subtle changes, caused by posttranslational modifications or amino acid variations, are manifested as significant changes in the activity of the protein's functions.

ACKNOWLEDGMENTS

We thank Ellis Golub for invaluable help in the three-dimensional protein analysis; Thomas Stamminger and Tom Shenk for reagents; and the members of the Alwine laboratory for helpful discussion and critical evaluation of the data. Cheers to all.

This work was supported by Public Health Service grant CA28379 awarded to J.C.A. by the National Cancer Institute.

REFERENCES

- Ahn, J.-H., Y. Xu, W.-J. Jang, M. J. Matunis, and G. S. Hayward. 2001. Evaluation of interactions of human cytomegalovirus immediate-early IE2 regulatory protein with small ubiquitin-like modifiers and their conjugation enzyme Ubc9. *J. Virol.* **75**:3859–3872.
- Arlt, H., D. Lang, S. Gebert, and T. Stamminger. 1994. Identification of binding sites for the 86-kilodalton IE2 protein of human cytomegalovirus within an IE2-responsive viral early promoter. *J. Virol.* **68**:4117–4125.
- Baracchini, E., E. Glezer, K. Fish, R. M. Stenberg, J. A. Nelson, and P. Ghazal. 1992. An isoform variant of the cytomegalovirus immediate-early auto repressor functions as a transcriptional activator. *Virology* **188**:518–529.
- Boyle, W. J., P. van der Geer, and T. Hunter. 1991. Phosphopeptide mapping and phosphoamino acid analysis by two-dimensional separation on thin-layer cellulose plates. *Methods Enzymol.* **201**:110–149.
- Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization and dynamics calculations. *J. Comput. Chem.* **4**:187–217.
- Caswell, R., L. Bryant, and J. Sinclair. 1996. Human cytomegalovirus immediate-early 2 (IE2) protein can transactivate the human *hsp70* promoter by alleviation of Dr1-mediated repression. *J. Virol.* **70**:4028–4037.
- Chiou, C.-J., J. Zong, I. Waheed, and G. S. Hayward. 1993. Identification and mapping of dimerization and DNA-binding domains in the C terminus of the IE2 regulatory protein of human cytomegalovirus. *J. Virol.* **67**:6201–6214.
- Damania, B., and J. C. Alwine. 1996. TAF-like function of SV40 large T antigen. *Genes Dev.* **10**:1369–1381.
- Ghazal, P., J. Young, E. Giulietti, C. DeMattei, J. Garcia, R. Gaynor, R. M. Stenberg, and J. A. Nelson. 1991. A discrete *cis* element in the human immunodeficiency virus long terminal repeat mediates synergistic *trans* activation by cytomegalovirus immediate-early proteins. *J. Virol.* **65**:6735–6742.
- Hagemeyer, C., S. Walker, R. Caswell, T. Kouzarides, and J. Sinclair. 1992. The human cytomegalovirus 80-kilodalton but not the 72-kilodalton immediate-early protein transactivates heterologous promoters in a TATA box-dependent mechanism and interacts directly with TFIID. *J. Virol.* **66**:4452–4456.
- Hagemeyer, C., S. M. Walker, P. J. G. Sissons, and J. H. Sinclair. 1992. The 72K IE1 and 80K IE2 proteins of human cytomegalovirus independently trans-activate the *c-fos*, *c-myc* and *hsp70* promoters via basal promoter elements. *J. Gen. Virol.* **73**:2385–2393.
- Harel, N. Y., and J. C. Alwine. 1998. Phosphorylation of the human cytomegalovirus 86-kilodalton immediate-early protein IE2. *J. Virol.* **72**:5481–5492.
- Heider, J. A., W. A. Bresnahan, and T. E. Shenk. 2002. Construction of a rationally designed human cytomegalovirus variant encoding a temperature-sensitive immediate-early 2 protein. *Proc. Natl. Acad. Sci. USA* **99**:3141–3146.
- Hermiston, T. W., C. L. Malone, P. R. Witte, and M. F. Stinski. 1987. Identification and characterization of the human cytomegalovirus immediate-early region 2 gene that stimulates gene expression from an inducible promoter. *J. Virol.* **61**:3214–3221.
- Hofmann, H., S. Floss, and T. Stamminger. 2000. Covalent modification of the transactivator protein IE2-p86 of human cytomegalovirus by conjugation to the ubiquitin-homologous proteins SUMO-1 and hSMT3b. *J. Virol.* **74**:2510–2524.
- Huang, L., C. L. Malone, and M. F. Stinski. 1994. A human cytomegalovirus early promoter with upstream negative and positive *cis*-acting elements: IE2 negates the effect of the negative element, and NF-Y binds to the positive element. *J. Virol.* **68**:2108–2117.
- Jones, D. T., W. R. Taylor, and J. M. Thornton. 1992. A new approach to protein fold recognition. *Nature* **358**:86–89.
- Just, M., A. Buerger-Wolff, G. Emoedi, and R. Hernandez. 1975. Immunisation trials with live attenuated cytomegalovirus TOWNE 125. *Infection* **3**:111–114.
- Kerry, J. A., M. A. Priddy, and R. M. Stenberg. 1994. Identification of sequence elements in the human cytomegalovirus DNA polymerase gene promoter required for activation by viral gene products. *J. Virol.* **68**:4167–4176.
- Klucher, K. M., M. Sommer, J. T. Kadonaga, and D. H. Spector. 1993. *In vivo* and *in vitro* analysis of transcriptional activation mediated by the human cytomegalovirus major immediate-early proteins. *Mol. Cell. Biol.* **13**:1238–1250.
- Leach, F. S., and E. S. Mocarski. 1989. Regulation of cytomegalovirus late-gene expression: differential use of three start sites in the transcriptional activation of ICP36 gene expression. *J. Virol.* **63**:1783–1791.
- Lukac, D. M., and J. C. Alwine. 1999. Effects of human cytomegalovirus major immediate-early proteins in controlling the cell cycle and inhibiting apoptosis: studies in ts13 cells. *J. Virol.* **73**:2825–2831.
- Lukac, D. M., N. Harel, and J. C. Alwine. 1997. TAF-like functions of the human cytomegalovirus immediate-early proteins. *J. Virol.* **71**:7227–7239.
- Lukac, D. M., J. R. Manuppello, and J. C. Alwine. 1994. Transcriptional activation by the human cytomegalovirus immediate-early proteins: requirements for simple promoter structures and interactions with multiple components of the transcription complex. *J. Virol.* **68**:5184–5193.
- Malone, C. L., D. H. Vesole, and M. F. Stinski. 1990. Transactivation of a human cytomegalovirus early promoter by gene products from the immediate-early gene IE2 and augmentation by IE1: mutational analysis of the viral proteins. *J. Virol.* **64**:1498–1506.
- Mocarski, E. S. 1996. Cytomegaloviruses and their replication, p. 2447–2492. *In B. N. Fields* (ed.), *Virology*. Raven Press, New York, N.Y.
- Monick, M. M., L. J. Geist, M. F. Stinski, and G. W. Hunninghake. 1992. The immediate early genes of human cytomegalovirus upregulate expression of the cellular genes *myc* and *fos*. *Am. J. Respir. Cell Mol. Biol.* **7**:251–256.
- Pizzorno, M. C., P. O'Hare, L. Sha, R. L. LaFemina, and G. S. Hayward. 1988. Transactivation and autoregulation of gene expression by the immediate-early region 2 gene products of human cytomegalovirus. *J. Virol.* **62**:1167–1179.
- Plotkin, S. A., T. Furukawa, N. Zygraich, and C. Huygelen. 1975. Candidate cytomegalovirus strain for human vaccination. *Infect. Immun.* **12**:521–527.
- Staprans, S. I., D. K. Rabert, and D. H. Spector. 1988. Identification of sequence requirements and *trans*-acting functions necessary for regulated expression of a human cytomegalovirus early gene. *J. Virol.* **62**:3463–3473.
- Stasiak, P. C., and E. S. Mocarski. 1992. Transactivation of the cytomegalovirus ICP36 gene promoter requires the beta gene product TRS1 in addition to IE1 and IE2. *J. Virol.* **66**:1050–1058.
- Stenberg, R. M. 1996. The human cytomegalovirus major immediate-early gene. *Intervirology* **39**:343–349.
- Stenberg, R. M., A. S. Depto, J. Fortney, and J. A. Nelson. 1989. Regulated expression of early and late RNAs and proteins from the human cytomegalovirus immediate-early gene region. *J. Virol.* **63**:2699–2708.
- Stenberg, R. M., J. Fortney, S. W. Barlow, B. P. Magrane, J. A. Nelson, and P. Ghazal. 1990. Promoter-specific *trans* activation and repression by human cytomegalovirus immediate-early proteins involves common and unique protein domains. *J. Virol.* **64**:1556–1565.
- Stenberg, R. M., and M. F. Stinski. 1985. Autoregulation of the human cytomegalovirus major immediate-early gene. *J. Virol.* **56**:676–682.
- Stenberg, R. M., P. R. Witte, and M. F. Stinski. 1985. Multiple spliced and unspliced transcripts from human cytomegalovirus immediate-early region 2 and evidence for a common initiation site within immediate-early region 1. *J. Virol.* **56**:665–675.
- Tanaka, M., V. Grossaikklaus, W. Herr, and N. Hernandez. 1988. Activation of the U2 SnRNA promoter by the octamer motif defines a new class of RNA polymerase II enhancer elements. *Genes Dev.* **2**:1764–1778.
- Waheed, I., C.-J. Chiou, J.-H. Ann, and G. S. Hayward. 1998. Binding of the HCMV 80-kDa immediate-early protein (IE2) to minor groove A/T rich sequences bounded by CG dinucleotides is regulated by protein oligomerization and phosphorylation. *Virology* **252**:235–257.
- Yoo, Y. D., C.-J. Chiou, K. S. Choi, S. Michelson, S. Kim, G. S. Hayward, and S.-J. Kim. 1996. The IE2 regulatory protein of human cytomegalovirus induces expression of the human transforming growth factor β 1 gene through an *egr-1* binding site. *J. Virol.* **70**:7062–7070.
- Yu, Y., and J. C. Alwine. 2002. Human cytomegalovirus major immediate-early proteins and simian virus 40 large T antigen can inhibit apoptosis through activation of the phosphatidylinositol 3'-OH kinase pathway and cellular kinase Akt. *J. Virol.* **76**:3731–3738.
- Yurochko, A. D., T. F. Kowalik, S.-M. Huong, and E.-S. Huang. 1995. Human cytomegalovirus upregulates NF- κ B activity by transactivating the NF- κ B p105/p50 and p65 promoters. *J. Virol.* **69**:5391–5400.
- Zhu, H., Y. Shen, and T. Shenk. 1995. Human cytomegalovirus IE1 and IE2 proteins block apoptosis. *J. Virol.* **69**:7960–7970.