JOURNAL OF VIROLOGY, Mar. 1997, p. 1857–1870 Vol. 71, No. 3
0022-538X/97/$04.00+0

Human Adenovirus Early Region 4 Open Reading Frame 1 Genes Encode Growth-Transforming Proteins That May Be Distantly Related to dUTP Pyrophosphatase Enzymes

ROBERT S. WEISS,[1] S. SYLVIA LEE,[1] B. V. VENKATARAM PRASAD,[1,2] AND RONALD T. JAVIER[1]*

Division of Molecular Virology[1] and Verna and Marrs Mclean Department of Biochemistry and W. M. Keck Center for Computational Biology,[2] Baylor College of Medicine, Houston, Texas 77030

Received 12 September 1996/Accepted 27 November 1996

An essential oncogenic determinant of subgroup D human adenovirus type 9 (Ad9), which uniquely elicits estrogen-dependent mammary tumors in rats, is encoded by early region 4 open reading frame 1 (E4 ORF1). Whereas Ad9 E4 ORF1 efficiently induces transformed foci on the established rat embryo fibroblast cell line CREF, the related subgroup A Ad12 and subgroup C Ad5 E4 ORF1s do not (R. T. Javier, J. Virol. 68:3917–3924, 1994). In this study, we found that the lack of transforming activity associated with non-subgroup D adenovirus E4 ORF1s in CREF cells correlated with significantly reduced protein levels compared to Ad9 E4 ORF1 in these cells. In the human cell line TE85, however, the non-subgroup D adenovirus E4 ORF1s produced protein levels higher than those seen in CREF cells as well as transforming activities similar to that of Ad9 E4 ORF1, suggesting that all adenovirus E4 ORF1 polypeptides possess comparable cellular growth-transforming activities. In addition, searches for known proteins related to these novel viral transforming proteins revealed that the E4 ORF1 proteins had weak sequence similarity, over the entire length of the E4 ORF1 polypeptides, with a variety of organismal and viral dUTP pyrophosphatase (dUTPase) enzymes. Even though adenovirus E4 ORF1 proteins lacked conserved protein motifs of dUTPase enzymes or detectable enzymatic activity, E4 ORF1 and dUTPase proteins were predicted to possess strikingly similar secondary structure arrangements. It was also established that an avian adenovirus protein, encoded within a genomic location analogous to that of the human adenovirus E4 ORF1s, was a genuine dUTPase enzyme. Although no functional similarity was found for the E4 ORF1 and dUTPase proteins, we propose that human adenovirus E4 ORF1 genes have evolved from an ancestral adenovirus dUTPase and, from this structural framework, developed novel transforming properties.

Human adenoviruses are organized into six different subgroups (A to F) and, in people, cause respiratory, gastrointestinal, and ocular infections (39, 40). Most human adenoviruses are also able to morphologically transform cultured rodent fibroblasts; in addition, subgroup A and B adenoviruses as well as subgroup D adenovirus type 9 (Ad9) are tumorigenic in rodents (28). Whereas subgroup A and B adenoviruses elicit undifferentiated sarcomas at the site of virus injection (28, 85), subcutaneous or intraperitoneal injection of subgroup D Ad9 into rats produces exclusively estrogen-dependent mammary tumors (2, 3, 42). For subgroup A and B adenoviruses, the viral E1 region, consisting of the E1A and E1B transcription units, is both necessary and sufficient for tumorigenicity (28). In contrast, subgroup D Ad9 requires early region 4 (E4) open reading frame 1 (ORF1) for its unique mammary oncogenicity (43, 45).

Analysis of available adenovirus DNA sequences indicates that all human adenoviruses, except those of subgroup F, code for E4 ORF1 proteins. Deletion of E4 ORF1 sequences does not measurably affect the replication of adenoviruses in standard permissive human cell lines, and for this reason, the function of E4 ORF1 in the viral life cycle has not been determined (8, 30, 41, 45). Nevertheless, in addition to its essential role in mammary tumor production by Ad9, the 14-kDa Ad9 E4 ORF1 (9ORF1) polypeptide alone is capable of inducing transformed foci, morphological alterations, anchorage-independent growth, increased saturation densities, and greatly enhanced oncogenicity in the established rat embryo fibroblast cell line CREF (45, 88). Together, these findings indicate that 9ORF1 encodes an oncoprotein. In CREF cells, the 9ORF1 protein is seen primarily within punctate cytoplasmic structures by indirect immunofluorescence (88). The mechanism by which this novel viral oncoprotein transforms cells remains to be determined.

Because subgroup D 9ORF1 is a potent transforming gene in CREF cells, we reasoned that related E4 ORF1s from other adenovirus subgroups would have similar abilities to transform cells. Consequently, it was surprising that expression plasmids encoding subgroup A Ad12 E4 ORF1 (12ORF1) and subgroup C Ad5 E4 ORF1 (5ORF1) fail to induce transformed foci on CREF cells (45). In the present study, we found that this seeming absence of transforming activity for the non-subgroup D virus E4 ORF1s was related to deficient protein expression in CREF cells. Moreover, using a human cell line that expressed higher levels of the non-subgroup D virus E4 ORF1 proteins, we were able to demonstrate that all E4 ORF1s had comparable growth-transforming activities. These results indicated that the adenovirus E4 ORF1s represent a family of functionally related transforming genes.

Additional findings reported in this study suggest that the adenovirus E4 ORF1 transforming proteins may be distantly related to dUTP pyrophosphatase (dUTPase) enzymes. dUTPases are essential cellular phosphoproteins (23, 27, 50) that serve two important purposes by catalyzing the hydrolysis of dUTP to dUMP and pyrophosphate. Reducing cellular dUTP pools decreases the potentially mutagenic misincorporation of uracil into replicating DNA (71); in addition, gener-

* Corresponding author. Mailing address: Division of Molecular Virology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030. Phone: (713) 798-3898. Fax: (713) 798-3586. E-mail: rjavier@bcm.tmc.edu.

ating dUMP yields a precursor for dTTP biosynthesis. A fundamental role for this enzyme in DNA replication is further underscored by the fact that human dUTPase exists in both nuclear and mitochondrial forms (51). Analogous to other cellular factors that participate in DNA replication, the levels of dUTPase enzymatic activity are cell cycle regulated and correlate with the proliferative state of cells (21, 34, 36, 65, 66, 81). Also relevant is the observation that dUTPases are frequently encoded by DNA viruses and retroviruses (22, 57). In this study, we determined that the adenovirus E4 ORF1 proteins exhibited sequence and predicted structural similarities with a variety of dUTPases, including a dUTPase encoded by an avian adenovirus. While a common function was not revealed, the results suggest that the adenovirus E4 ORF1 transforming proteins may be structurally and evolutionarily related to an important enzyme of nucleotide metabolism.

MATERIALS AND METHODS

Cell lines, PCR primers, and plasmids. CREF (26) and TE85 (56) cell lines were grown in culture medium (Dulbecco's modified Eagle medium supplemented with 20 μg of gentamicin per ml) containing 6 and 10% fetal bovine serum (FBS), respectively. Cells were maintained in a humidified incubator under a 5% CO_2 atmosphere at 37°C.

Subgroup A 12ORF1, subgroup C 5ORF1, and subgroup D 9ORF1 were PCR amplified as previously described (45). Subgroup B Ad3 E4 ORF1 (3ORF1), avian adenovirus type 1 CELO dUTPase, and human dUTPase were PCR amplified from Ad3 viral DNA, CELO viral DNA, and plasmid pHUMDUT1 (58) templates, respectively, using the following primer sets flanked by *Bam*HI and *Eco*RI sites (restriction enzyme sites underlined): 3ORF1, 5′CTC GGA TCC ATG ATT GAG GCT TTG TAT3′ and 5′CTC GAA TTC AAA CAA GCG TCA TAT CAT3′; CELO dUTPase, 5′CTC GGA TCC ATG GAC CCG TTC GGT TCT3′ and 5′CTC GAA TTC TCA ACG GGA ACC AGG GGT3′; and human dUTPase, 5′CTC GGA TCC ATG CAG CTC CGC TTT GCC3′ and 5′CTC GAA TTC TTA ATT CTT TCC AGT GGA3′. The primer set for 3ORF1 was based on the related Ad34 E4 DNA sequences (15). E4 ORF1 and dUTPase PCR products yielded full-length polypeptides except for the human dUTPase PCR product, for which the published amino acid sequence (58), coding for an enzymatically active protein (18), was used but has recently been demonstrated to be missing the amino-terminal 23 residues of the wild-type nuclear enzyme (51). The influenza virus hemagglutinin (HA) epitope tag (24) was placed at the amino terminus of each E4 ORF1 from adenovirus subgroups A to D by PCR amplification from viral DNA templates, using the following primer sets (HA epitope tags underlined): HA-12ORF1, 5′ATG GCT TAC CCA TAC GAT GTA CCT GAC TAT GCG GCT GCT TTT GAG ACT CTT3′ and 5′CTC GAA TTC TTA AAT TAA AGA TGC CTG3′; HA-3ORF1, 5′ATG GCT TAC CCA TAC GAT GTA CCT GAC TAT GCG ATT GAG GCT TTG TAT GTC3′ and 5′CTC GAA TTC AAA CAA GCG TCA TAT CAT3′; HA-5ORF1, 5′ATG GCT TAC CCA TAC GAT GTA CCT GAC TAT GCG GCT GCC GCT GTG GAA GCG3′ and 5′CTC GAA TTC TTA AAC ATT AGA AGC CTG3′; and HA-9ORF1, 5′ATG GCT TAC CCA TAC GAT GTA CCT GAC TAT GCG GCT GAA TCT CTG TAT GCT3′ and 5′CTC GAA TTC ACA GAA AGC ATG CAC TAA3′.

For the construction of cytomegalovirus (CMV) expression plasmids, wild-type and tagged E4 ORF1 PCR products were blunt ended and cloned into the blunt-ended *Bam*HI site of the pCMV-Neo-Bam_3 expression plasmid (32). Wild-type 12ORF1, 5ORF1, and 9ORF1 CMV expression plasmids were described previously (45). For the construction of plasmids encoding glutathione *S*-transferase (GST) fusion proteins, wild-type E4 ORF1 and dUTPase PCR products were digested with *Bam*HI and *Eco*RI and cloned into the corresponding restriction enzyme sites of plasmid pGEX-2T (Pharmacia Biotech Inc.). All plasmid constructs were verified by sequencing.

Cell pools of stable transformants and transformation assays. Cell pools of stable CREF and TE85 transformants were generated as previously described (88). Briefly, 20 μg of plasmid DNA was transfected onto subconfluent cells by calcium phosphate precipitation with a glycerol shock (47). At 72 h posttransfection, cells were passaged at low density (5×10^5 cells per 10-cm-diameter dish) into culture medium supplemented with FBS and G418 (Gibco BRL Life Technologies, Inc.). Selections with CREF and TE85 cells used G418 concentrations of 360 and 115 μg/ml, respectively. Each cell pool was produced by collecting all G418-resistant colonies (more than 100) appearing on a single 10-cm-diameter dish after 10 days of selection in G418. Focus and soft-agar transformation assays were performed as previously described (45, 88).

Immunoblot analyses. Cell lysates were prepared in radioimmunoprecipitation assay buffer (50 mM Tris-HCl [pH 8.0], 150 mM NaCl, 1% [vol/vol] Nonidet P-40, 0.5% [wt/vol] sodium deoxycholate, 0.1% [wt/vol] sodium dodecyl sulfate) containing protease inhibitors (300 μg of phenylmethylsulfonyl fluoride per ml and 6 μg each of aprotinin and leupeptin per ml). Immunoblot analyses were performed as described previously (88), using the HA epitope-specific monoclonal antibody 12CA5 (Berkeley Antibody Co.).

GST fusion proteins and dUTPase enzyme assays. The purification of GST fusion proteins from bacteria transformed with pGEX-2T constructs was performed by standard methods (76). Isopropyl-β-D-galactoside (IPTG)-induced bacteria were lysed in Triton X-100 solubilization buffer (50 mM Tris-HCl [pH 8.0], 120 mM NaCl, 50 mM EDTA, 10 mM β-mercaptoethanol, 3 mg of lysozyme per ml, 1% [vol/vol] Triton X-100, protease inhibitors) and centrifuged at $12,000 \times g$ for 15 min, and the resulting bacterial supernatants were incubated with glutathione-Sepharose beads to isolate soluble GST proteins. After several washes of the beads with NTN buffer (20 mM Tris-HCl [pH 8.0], 100 mM NaCl, 0.5% [vol/vol] Nonidet P-40), GST proteins were eluted into elution buffer (0.1 M Tris-HCl [pH 8.0], 120 mM NaCl, 40 mM glutathione), examined on a protein gel with standards to ascertain their concentration and purity, and adjusted to a final concentration of 1 ng per μl with elution buffer for subsequent use in dUTPase enzyme assays (see below). As a negative control standard for the enzyme assays, mock-protein preparations were prepared identically to the GST proteins, using bacterial supernatants derived from IPTG-induced bacteria transformed by plasmid pUC19.

dUTPase enzyme assays were performed as described by others (48, 89). An equal volume of either purified GST protein, elution buffer (no protein), or mock-protein preparation, brought to 50 μl with elution buffer, was mixed with 50 μl of reaction buffer (0.1 M Tris-HCl [pH 8.0], 4 mM β-mercaptoethanol, 2 mM $MgCl_2$, 4 mM *p*-nitrophenylphosphate, 0.2% bovine serum albumin, 1 μCi of deoxy[5-^{3}H]uridine 5′-triphosphate [15 Ci/mmol]) and incubated for 5 min at 37°C. Half of each reaction mixture was spotted onto a 2.3-cm-diameter DE-81 paper disc, which was washed three times with 1 mM ammonium formate–4 M formic acid, washed once with 95% ethanol, and dried. The amount of radioactivity retained on the paper disc was quantitated with a liquid scintillation counter. Data are reported as percent dUTP hydrolyzed by each protein relative to the amount of dUTP remaining following an enzyme reaction with the mock-protein preparation, which routinely yielded the lowest dUTP hydrolysis values. For dUTPase enzyme substrate competition experiments, 100 μM unlabeled dUTP, UTP, or dATP was included in the reaction mixtures, which were processed as described above.

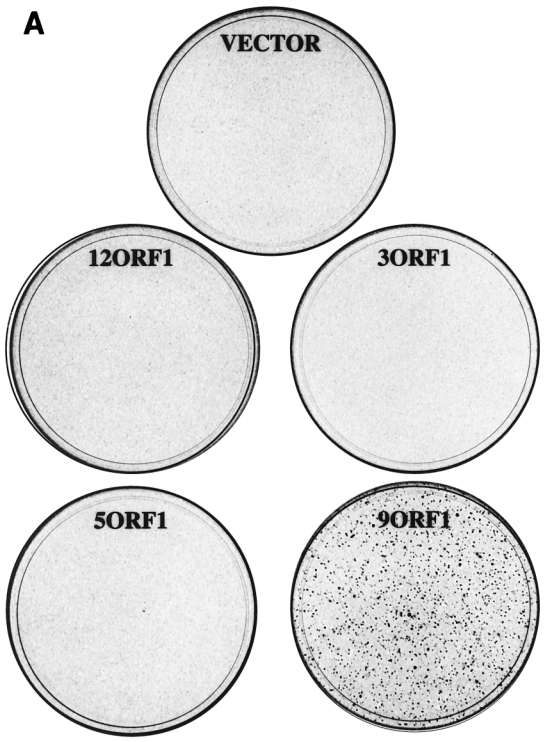
Protein sequences, alignments, secondary structure predictions, and phylogenetic analyses. Sequences used in this study were derived from GenBank. Human adenovirus E4 ORF1 and porcine adenovirus type 3 (PAV-3) E4 ORF3 protein sequences were used intact for the sequence analyses, whereas additional nonconserved amino- or carboxy-terminal sequences for some dUTPases were trimmed to produce dUTPase core sequences (see Fig. 5). The dUTPase core sequences were used in dUTPase:dUTPase pairwise sequence comparisons, multiple sequence analyses, and secondary structure predictions. For E4 ORF1:dUTPase pairwise sequence comparisons and phylogenetic analyses, dUTPases were further trimmed to generate truncated dUTPase core sequences, which represented the regions of the dUTPase polypeptides having sequence similarity with the E4 ORF1 proteins (see Fig. 4).

Pairwise and multiple sequence alignments were performed with the GAP algorithm of the Genomics Computing Group, Inc. (GCG), software package and the PIMA software program version 1.4, respectively (77, 78). Protein secondary structures were predicted by the PHD secondary structure prediction program (Heidelberg University) (73), and phylogenetic trees were generated by the parsimony method, using the PAUP software program version 3.1 (83).

RESULTS

Non-subgroup D adenovirus E4 ORF1s exhibit weak transforming activity compared to 9ORF1 in CREF cells. Using CMV expression plasmids, we previously reported that subgroup D 9ORF1 potently induces transformed foci on CREF cells, whereas subgroup A 12ORF1 and subgroup C 5ORF1 do not (45). In addition to confirming these previous results, we found that a subgroup B 3ORF1 CMV expression plasmid also fails to generate transformed foci on CREF cells (Fig. 1A), indicating that among the representative E4 ORF1s from each adenovirus subgroup, only subgroup D 9ORF1 was capable of focus formation on these cells.

The large difference in focus-forming potential on CREF cells between 9ORF1 and non-subgroup D virus E4 ORF1s was unexpected for several reasons. The E4 ORF1 genes are situated at identical positions within the genomes of subgroup A to D human adenoviruses (15, 31, 33, 44), and the amino acid sequences of their predicted polypeptides showed significant similarity, ranging from 44.8 to 51.2% amino acid identity and 64.8 to 69.4% amino acid similarity (Table 1). This degree of protein similarity was comparable to that of subgroup A to



B

```
12ORF1    M-AAFETLYVYFTGPGAMLPKQEGDSNAYVLFSPANFVIPPHGVV   44
3ORF1     M---IEALYVYLEGPGATLPEQQQRNN-YIFYSPVPFTLYPRGVA   41
5ORF1     MAAAVEALYVVLEREGAILPRQEGFSGVYVFFSPINFVIPPMGAV   45
9ORF1     M---AESLYAFIDSPGGIAPVQEGTSNRYTFFCPESFHIPPHGVV   42

12ORF1    LLYLHIAVDIPPGYLGTLFSLCDMNARGVFVGAETLYPGSRMELS   89
3ORF1     LLYLKLSIIIPRGYVGCFFSLTDANMSGLYASSRIIHAGHREELS   86
5ORF1     MLSLRLRVCIPPGYFGRFLALTDVNQPDVFTESYIMTPDMTEELS   90
9ORF1     LLHLKVSVLVPTGYQGRFMALNDYHARDILTQSDVIFAGRRQELT   87

12ORF1    VLLFNHSCVFCDVRAKQPVARLLLSRVVFPPVCQASLI   127
3ORF1     VLLFNHNDRFYEGRAGDPVACLVMERLIFPPVRQATMI   124
5ORF1     VVLFNHGDQFFYGHAGMAVVRLMLIRVVFPVVRQASNV   128
9ORF1     VLLFNHTDRFLYVRKGHPVGTLLLERVIFPSVKIATLV   125
```

FIG. 1. (A) Focus transformation assays in CREF cells with wild-type human adenovirus E4 ORF1s. Focus assays were performed as described previously (45) by transfecting subconfluent CREF cells on 10-cm-diameter dishes with 20 μg of either the empty CMV expression plasmid (vector) or CMV expression plasmids encoding the indicated adenovirus E4 ORF1s. After 4 weeks, transformed foci were visualized by fixing the cells in methanol and staining them with Giemsa stain. (B) Sequence alignments of the human adenovirus E4 ORF1 proteins from subgroups A to D. The multiple sequence alignment was performed with the PIMA multiple sequence alignment algorithm (77, 78). Boldface letters designate amino acid residues conserved between the four human adenovirus E4 ORF1 proteins.

D and F adenovirus 13S E1A proteins, which exhibited 36.8 to 50.2% amino acid identity and 60.3 to 69.7% amino acid similarity (reference 74 and data not shown). Also, a multiple protein sequence alignment further revealed that greater than 25% of the E4 ORF1 amino acid residues were conserved for all of the polypeptides (Fig. 1B). Together, these results were consistent with the idea that all human adenovirus E4 ORF1 proteins are both evolutionarily and functionally related and suggested that, despite the results of focus assays with CREF cells, all adenovirus E4 ORF1s might possess at least some cell growth-transforming activity. Therefore, with this expectation, we performed additional experiments in an attempt to expose transforming activities for the nonsubgroup D virus E4 ORF1s.

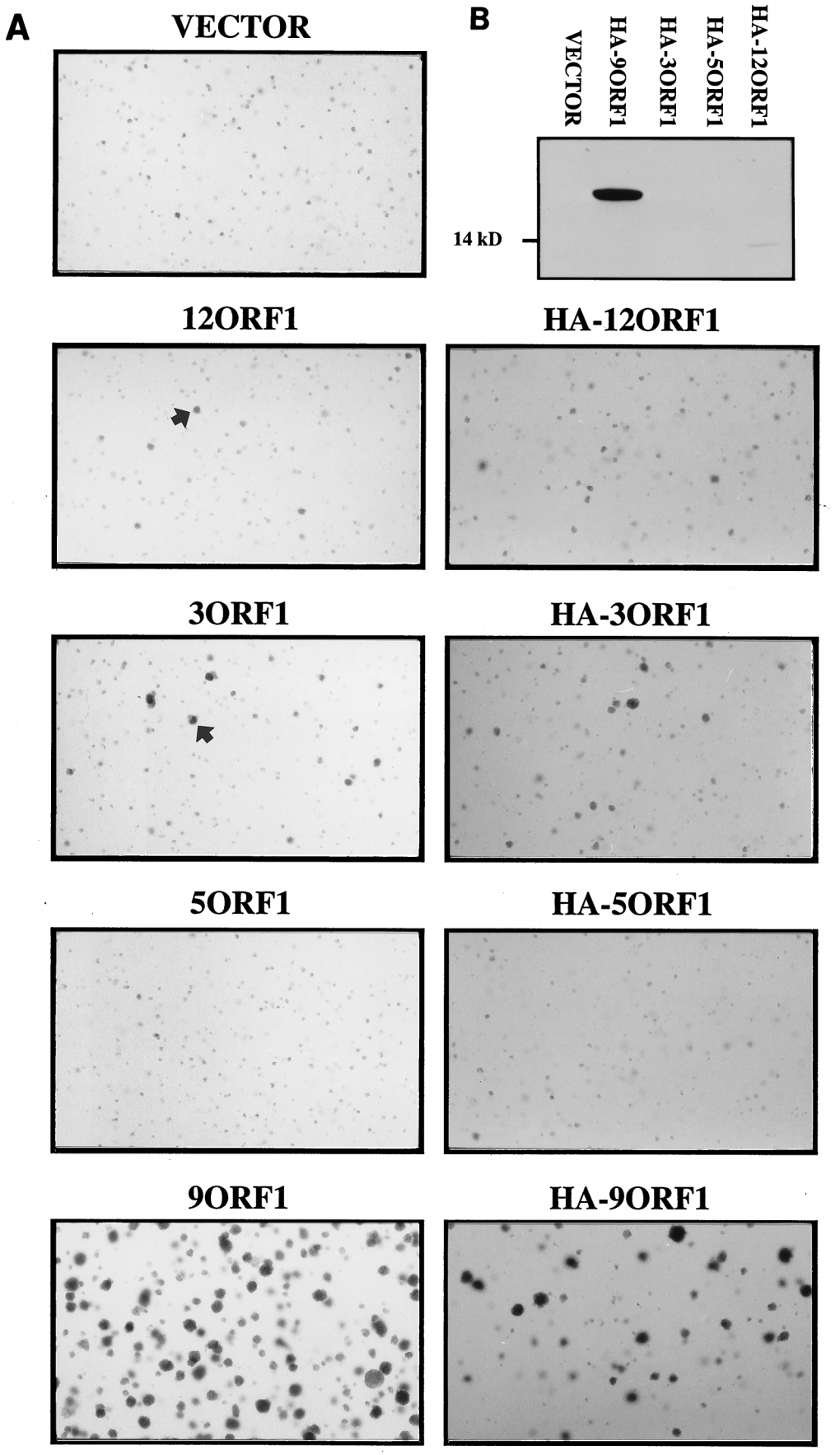
In separate studies with transformation-defective mutant 9ORF1 proteins, we have observed that growth in soft agar is a more sensitive assay than focus formation for detecting weak transforming activity with CREF cells (unpublished results). Consequently, we reasoned that reduced transforming activities for the non-subgroup D virus E4 ORF1s might be observed in soft-agar transformation assays. Using an empty CMV expression plasmid or CMV plasmids encoding each E4 ORF1, we established cell pools of stable CREF cell transformants by G418 selection and tested the resulting cells for the ability to form colonies in soft agar. During these selections, the non-subgroup D virus E4 ORF1 plasmids yielded equivalent numbers of G418-resistant colonies as the 9ORF1 plasmid, suggesting that the lack of transforming activity for the former plasmids was not the result of toxicity (data not shown). Consistent with the focus assays shown in Fig. 1A, the transforming activities in these soft agar assays for the non-subgroup D virus E4 ORF1s were significantly lower than that of 9ORF1 (Fig. 2A), which, as reported previously (88), potently induced anchorage-independent growth in CREF cells (cloning efficiency of approximately 30%). Nevertheless, whereas the 5ORF1 plasmid failed to elicit any detectable transforming activity, the 12ORF1 and 3ORF1 plasmids consistently induced small numbers of CREF cell colonies (cloning efficiencies of approximately 1 to 3%) larger than those produced by the empty CMV plasmid. These results gave the first indication that, besides 9ORF1, other E4 ORF1s may indeed possess transforming activity.

Non-subgroup D adenovirus E4 ORF1s display significantly lower protein levels than 9ORF1 in CREF cells. Lower protein levels for the non-subgroup D virus E4 ORF1s compared to 9ORF1 might account for their weaker transforming activities in CREF cells. Consequently, we next investigated E4 ORF1 protein expression in these cells. For these analyses, we chose to tag the E4 ORF1 proteins at their amino termini with an HA epitope because this approach would allow us to directly compare their steady-state protein levels in cells. The tag did not significantly alter the transforming phenotypes of the E4 ORF1s, as CREF cell pools selected for each corresponding pair of wild-type and tagged E4 ORF1 CMV plasmids yielded similar results in soft-agar transformation assays (Fig. 2A). We next compared protein expression of the different tagged E4 ORF1s in these same CREF cell pools by immunoblot analysis (Fig. 2B). The results of these analyses indicated that non-subgroup D virus E4 ORF1 proteins were expressed at significantly lower levels than the 9ORF1 protein, with only small amounts of 12ORF1 protein and no 3ORF1 or 5ORF1 protein being detected. In protein gels, the tagged 12ORF1 displayed a mobility consistent with its predicted molecular weight (Fig. 2B), as did the other tagged non-subgroup D virus E4 ORF1

TABLE 1. Sequence similarities between subgroup A to D human adenovirus E4 ORF1 proteins

Protein	% Amino acid identity[a] (% amino acid similarity) to:				
	12ORF1	3ORF1	5ORF1	9ORF1	PAV-3 E4 ORF3
PAV-3 E4 ORF3	19.4 (48.4)	22.4 (47.4)	21.0 (45.2)	26.6 (47.6)	100
9ORF1	48.8 (67.2)	45.2 (67.7)	44.8 (64.8)	100	
5ORF1	51.2 (67.7)	46.0 (66.1)	100		
3ORF1	48.4 (69.4)	100			
12ORF1	100				

[a] Determined for full-length sequences by using the GAP algorithm (gap creation penalty = 3.00; gap extension penalty = 0.10) of the GCG sequence analysis software package. PAV-3 E4 ORF3 was included because it may be an E4 ORF1 homolog (see text).


A
VECTOR
B
VECTOR
HA-9ORF1
HA-3ORF1
HA-5ORF1
HA-12ORF1
14 kD
12ORF1
HA-12ORF1
3ORF1
HA-3ORF1
5ORF1
HA-5ORF1
9ORF1
HA-9ORF1

proteins (see below), but for unknown reasons, the tagged 9ORF1 protein migrated more slowly. Reduced protein expression by the tagged non-subgroup D virus E4 ORF1s compared to tagged 9ORF1 was also seen following transient transfection of plasmids in CREF cells (data not shown). Moreover, the same transformation and protein expression results as those described for the E4 ORF1s in CREF cells were obtained in the murine fibroblast cell line C127 as well (data not shown). While functional defects for the non-subgroup D virus E4 ORF1 proteins in CREF and C127 cell lines cannot be ruled out, insufficient protein expression was the simplest explanation for their lower transforming potentials compared to 9ORF1.

Expression of higher levels of the non-subgroup D adenovirus E4 ORF1 proteins in the human TE85 cell line results in transforming activity comparable to that of 9ORF1. One possible inference from the results presented above was that higher expression of the non-subgroup D virus E4 ORF1 proteins in cells would produce transforming activities like those of the 9ORF1 protein. To test this hypothesis, we sought to identify a nontransformed cell line that could express non-subgroup D virus E4 ORF1 proteins at levels equivalent to that of the 9ORF1 protein. Because low expression of non-subgroup D virus E4 ORF1 proteins had been observed in two different rodent fibroblast cell lines (CREF and C127) and because human cells represent the natural host for adenovirus infections, we reasoned that expression of all E4 ORF1 proteins might be the same in a human cell line. In support of this notion, human HeLa and 293 cell lines expressed comparable levels of all E4 ORF1 proteins (data not shown). These highly transformed cells, however, are unsuitable for transformation studies. These observations led us to evaluate human TE85 cells, a nontumorigenic anchorage-dependent cell line derived from an osteosarcoma (56).

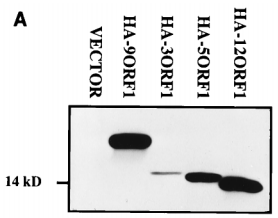
As with the CREF cell analyses described above, cell pools of stable TE85 transformants were selected for the empty CMV plasmid and CMV plasmids containing each tagged E4 ORF1 and analyzed for protein expression by immunoblot analysis. From these experiments, it was determined that human TE85 cells expressed most non-subgroup D virus E4 ORF1 proteins at levels comparable to that of the 9ORF1 protein, with tagged 3ORF1 protein being expressed at lower levels (Fig. 3A). With respect to the findings with CREF cells (Fig. 2B), these results also demonstrated that the non-subgroup D E4 ORF1 expression plasmids were not generally defective for protein expression. More important, for these TE85 cell pools expressing higher relative amounts of the non-subgroup D virus E4 ORF1 proteins than CREF cells, all of the E4 ORF1 proteins, including the 3ORF1 protein, conferred similar transforming activities in soft-agar transformation assays (Fig. 3B). The mean cloning efficiencies for the TE85 cells expressing the tagged E4 ORF1s in these experiments were 17.8% ± 2.0% (12ORF1), 30.9% ± 4.2% (3ORF1), 42.0% ± 4.3% (5ORF1), and 27.6% ± 3.1% (9ORF1). Similar results were also obtained with TE85 cell pools selected for the wild-type E4 ORF1 expression plasmids (Fig. 3B). Significantly, these findings suggested that all human adenovirus E4 ORF1 proteins exhibit comparable growth-transforming activities when assayed in an appropriate cell line.

Adenovirus E4 ORF1 proteins have limited sequence similarity with dUTPase enzymes. To identify proteins that might be related to the E4 ORF1 transforming proteins, we performed BLAST searches of databases with these viral sequences. Besides the expected similarities between the E4 ORF1 proteins themselves (Table 1), these searches also revealed weaker sequence similarities to an assortment of dUTPase enzymes. The known or putative dUTPases showing sequence similarity to all E4 ORF1 proteins were derived from humans (*Homo sapiens*), yeasts (*Candida albicans* and *Saccharomyces cerevisiae*), plants (*Lycopersicon esculentum* [tomato]), and two types of DNA viruses (adenovirus [avian adenovirus CELO] and poxviruses [vaccinia, variola, and Orf viruses]). In some cases, similarities with retrovirus (e.g., feline immunodeficiency virus, equine infectious anemia virus [EIAV], and visna virus) dUTPases were also detected. In addition to the sequence similarities observed between E4 ORF1 and the dUTPase proteins, it was also significant that one putative dUTPase was encoded by an avian adenovirus distantly related to human adenoviruses (16).

To determine whether E4 ORF1 and dUTPase sequence similarities were limited to a portion or the entire length of either protein, we performed pairwise alignments between full-length E4 ORF1 and dUTPase sequences. From such analyses, it was determined that the sequence similarities between E4 ORF1 and dUTPase proteins extended over the entire lengths of the E4 ORF1 proteins and that dUTPases were longer than E4 ORF1 proteins by invariably having extended carboxy termini and by sometimes also having additional sequences at their amino termini. The nonconserved extended amino-terminal residues of human dUTPase, and probably other dUTPases as well, are dispensable for enzymatic activity (18). Typical sequence similarities and differences observed between the E4 ORF1 and dUTPase polypeptides are illustrated by two representative pairwise alignments shown in Fig. 4.

We next quantitated sequence similarities between the E4 ORF1 proteins and 18 different dUTPase proteins, including several from prokaryotes. Because dUTPases are longer than E4 ORF1 proteins, dUTPases were trimmed to those sequences having similarity to the E4 ORF1 sequences for these analyses (see Fig. 4 and Materials and Methods). Using these truncated dUTPase core sequences, we tabulated protein similarity values (percent amino acid identity and similarity) for all possible combinations of pairwise alignments with the full-length E4 ORF1 sequences (Table 2). The results indicated that the dUTPases from *C. albicans*, *S. cerevisiae*, and *L. esculentum*, as well as avian adenovirus CELO, showed the highest similarity values by being the only enzymes having ≥21.0% identity and ≥49.6% similarity with all of the E4 ORF1 proteins. In general, however, the similarity values between E4 ORF1 and dUTPase proteins were low, ranging from 9.6 to 29.8% identity and 36.7 to 57.3% similarity. Nevertheless, some of the highest similarity values between the E4 ORF1

FIG. 2. (A) Soft-agar transformation assays in CREF cells with wild-type and epitope-tagged human adenovirus E4 ORF1s. CREF cell pools of stable transformants were generated as described in Materials and Methods with the empty CMV expression plasmid (vector) or CMV expression plasmids encoding the indicated wild-type or HA epitope-tagged adenovirus E4 ORF1s. Soft-agar assays were performed by suspending 10^5 viable cells in culture medium containing 10% FBS and 0.4% Noble agar (88). Photographs were taken 18 days (vector and wild-type E4 ORF1s) or 21 days (epitope-tagged E4 ORF1s) after plating in soft agar (magnification, ×18.6). Arrows indicate examples of 12ORF1- and 3ORF1-induced colonies larger than those produced by the empty CMV plasmid. (B) Expression of HA epitope-tagged E4 ORF1 proteins in CREF cell pools. CREF cells from the pools used for panel A were lysed in radioimmunoprecipitation assay buffer, and 150 μg of each of the resulting proteins was resolved on a 15% Laemmli gel, transferred to a polyvinylidene difluoride membrane, and immunoblotted with the HA-specific monoclonal antibody 12CA5.


A
VECTOR
HA-9ORF1
HA-3ORF1
HA-5ORF1
HA-12ORF1
14 kD

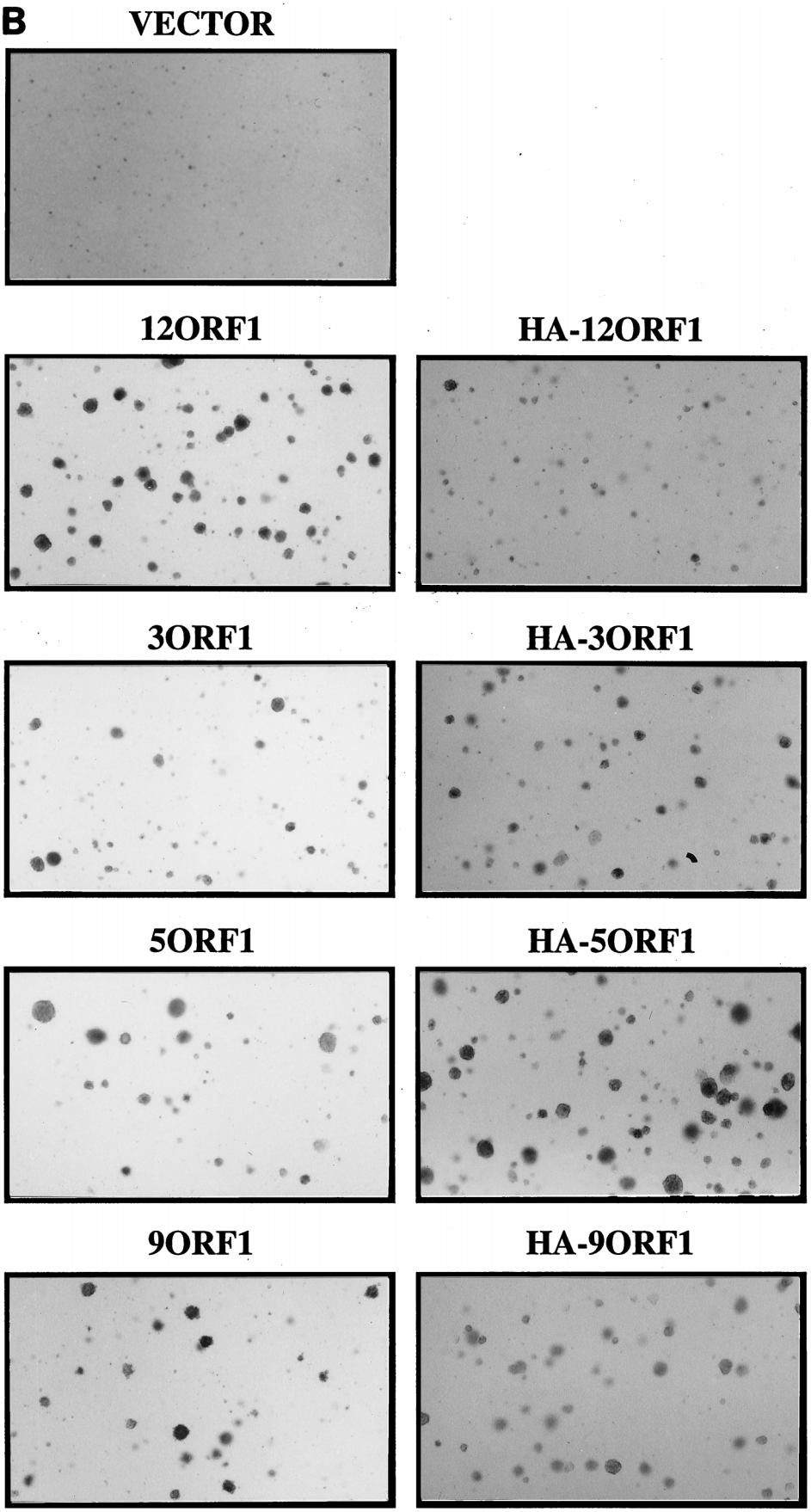

B
VECTOR
12ORF1
HA-12ORF1
3ORF1
HA-3ORF1
5ORF1
HA-5ORF1
9ORF1
HA-9ORF1

FIG. 3. (A) Expression of HA epitope-tagged human adenovirus E4 ORF1s in TE85 cells. TE85 cell pools of stable transformants were generated as described in Materials and Methods with either the empty CMV expression plasmid (vector) or CMV expression plasmids encoding the indicated HA epitope-tagged adenovirus E4 ORF1s. Expression of HA epitope-tagged E4 ORF1 proteins was determined by immunoblot analysis as described in the legend to Fig. 2. (B) Soft-agar transformation assays in TE85 cells with wild-type and epitope-tagged human adenovirus E4 ORF1s. TE85 cell pools of stable transformants were generated as described for panel A, and soft-agar assays were performed as described in the legend to Fig. 2. Photographs were taken 16 days (vector and epitope-tagged E4 ORF1s) or 21 days (wild-type E4 ORF1s) after plating in soft agar (magnification, ×18.6).

and dUTPase proteins were higher than the lowest similarity values between the dUTPase core sequences (see Materials and Methods) of distantly related dUTPase proteins (Table 3). For example, the E4 ORF1:dUTPase similarity values between 9ORF1:*C. albicans*, 3ORF1:*C. albicans*, and 3ORF1:CELO (see Table 2, boxes) were all higher than the dUTPase:dUTPase similarity values between *Mycobacterium leprae*:EIAV, Shope fibroma virus:EIAV, CELO:EIAV, Shope fibroma virus:visna virus, and Orf virus:visna virus (Table 3, boxes). In summary, the results presented above suggested that E4 ORF1 and dUTPase proteins may be distantly related.

Adenovirus E4 ORF1 proteins are not dUTPase enzymes. In light of the observed sequence similarities, we wanted to determine whether E4 ORF1 proteins had dUTPase enzymatic activity. The likelihood of this possibility was first assessed by further comparisons of the E4 ORF1 and dUTPase protein sequences. Known dUTPase enzymes display five conserved motifs (I to V) which are believed to function in substrate binding and catalysis (52, 57). In particular, motif V exhibits similarity with the phosphate-binding-loop domains of nucleotide-binding proteins (13, 57, 62) and hence has been postulated to participate in dUTPase binding to the phosphate residues of dUTP. We hypothesized that if E4 ORF1 proteins were to have dUTPase enzymatic activity, the conserved dUTPase domains would be maintained in the E4 ORF1 protein sequences. To facilitate the identification of possible dUTPase motifs within the E4 ORF1 proteins, we aligned the E4 ORF1 sequences and dUTPase core sequences from the 18 dUTPase proteins and, using the locations of the dUTPase motifs as guides, looked for conserved amino acid residues within analogous positions of the adenovirus E4 ORF1 polypeptides (Fig. 5). The results clearly demonstrated that dUTPase motif V was missing in the E4 ORF1 proteins, as it resided within the carboxy-terminal dUTPase sequences absent from the E4 ORF1 proteins. In addition, close inspection of E4 ORF1 sequences in regions equivalent to the remaining four conserved dUTPase motifs indicated that the E4 ORF1 proteins preserved few, if any, of the conserved amino acid residues within these locations. The results of these sequence comparisons suggested that E4 ORF1 proteins do not have dUTPase enzymatic activity.

We next tested this prediction by assaying the E4 ORF1 proteins for dUTPase enzymatic activity. For this purpose, we used purified recombinant GST fusion proteins expressed in bacteria because, for several different dUTPases, this approach yields fusion proteins with enzymatic activity (10, 18, 48). A GST-human dUTPase fusion protein, similar to an enzymatically active recombinant protein reported by others (18), was used as a positive control in these experiments. In the dUTPase enzyme assays, this GST-human dUTPase construct exhibited high levels of enzymatic activity (92.1% dUTP hydrolysis) compared to the negative control GST protein (12.2% dUTP hydrolysis) (Fig. 6A). Moreover, consistent with the E4 ORF1 proteins not preserving dUTPase conserved domains, none of the GST-E4 ORF1 proteins showed dUTPase activity above that of the GST negative control. The absence of de-

A

```
                    .         .         .         .         .
human dUTPase    1 MPCSEETPAISPSKRARPAEVGGMQLRFARLSEHATAPTRGSARAAGYDL 50
                                        :  :| :...:. :. ..: .. |.:
9ORF1            1                      MAESLYAFIDSPGGIAPVQEGTSNRYTF 28

                    .         .         .         .         .
human dUTPase   51 YSAYDYTIPPMEKAVVKTDIQIALPSGCYGRVAPRSGLAAKHFIDVGAGV 100
                   ::: .: |:| : .::. .: : :|.|: ||. : .:. |:.:: ...:|
9ORF1           29 FCPESFHIPPHGVVLLHLKVSVLVPTGYQGRFMALNDYHARDIL.TQSDV 77

                    .         .         .         .         .         .
human dUTPase  101 IDEDYRGNVGVVLFNFGKEKFEVKKGDRIAQLICERIFYPEIEEVQALDDTERGSGGFGSTGKNS 164
                   | .: |.::.|:||| ..  : |:||..:: |: ||:::|.:. .  :
9ORF1           78 IFAGRRQELTVLLFNHTDRFLYVRKGHPVGTLLLERVIFPSVKIATLV                  125

                    .         .         .         .         .
Shope dUTPase   1   MSLYVKCVKLSNNAIIPNRSMSGSAGYDLYSAYSYTVKPYNRILVRTD 48
                      .: . :| |..:: |  |  : |:.| ::|:..:.:.| . :::.
5ORF1           1 MAAAVEALYVVLEREGAILPRQEGFSGVYVFFSPINFVIPPMGAVMLSLR 50

                    .         .         .         .         .
Shope dUTPase  49 ICLMIPDKCYGRISPRSGLSLNYNIDIGGGVIDSDYRGEIGIVFINNGCS 98
                  : : ||. ::||: : .::.   :: .:: ::..|  :|:::|::|:|
5ORF1          51 LRVCIPPGYFGRFLALTDVNQP.DVFTESYIMTPDMTEELSVVLFNHGDQ 99

                    .         .         .         .
Shope dUTPase  99 DFNIKVGDRIAQIIFERVEYPIMEEVKCLEDTERGNSGFGSSGM 142
                    |  ..|  :..::: || :|:: :.. :
5ORF1         100 FFYGHAGMAVVRLMLIRVVFPVVRQASNV               128
```

B

FIG. 4. Pairwise alignments between representative full-length adenovirus E4 ORF1 and dUTPase protein sequences. (A) Human dUTPase and 9ORF1 protein sequences. (B) Shope fibroma virus dUTPase and 5ORF1 protein sequences. Pairwise alignments were generated by using the GAP algorithm of the GCG software package (gap creation penalty = 3.00; gap extension penalty = 0.10). Positions of amino acid identity are indicated by vertical lines, whereas high and low degrees of similarity between amino acid residues are denoted by double and single dots, respectively. For the two representative alignments, note that (i) the sequence similarity extends over the entire length of the E4 ORF1 protein, (ii) human dUTPase has additional amino- and carboxy-terminal amino acid residues compared to the E4 ORF1 protein, and (iii) Shope fibroma virus dUTPase has only additional carboxy-terminal sequences compared to the E4 ORF1 protein. The five conserved motifs of dUTPase enzymes are indicated in Fig. 5.

TABLE 2. Sequence similarities between adenovirus E4 ORF1 proteins and 18 different dUTPase proteins

Protein	% Amino acid identity[a] (% amino acid similarity) to dUTPase protein from:																
	Prokaryote			Eukaryote			DNA virus					Retrovirus					
	E. coli	*H. influenzae*	*C. burnetii*	*M. leprae*	*C. albicans*	*S. cerevisiae*	*H. sapiens*	*L. esculentum*	Vaccinia virus	Shope fibroma virus	HSV-1	Orf virus	Avian adenovirus CELO	Visna virus	SRV-1	MMTV	EIAV	FIV
12ORF1	22.7 (52.2)	18.9 (44.3)	11.0 (46.5)	27.6 (52.8)	27.8 (56.3)	22.4 (50.4)	19.5 (53.7)	24.6 (54.8)	19.0 (50.8)	16.0 (52.8)	20.0 (49.6)	23.0 (53.2)	24.6 (56.3)	14.2 (45.8)	17.0 (39.0)	18.7 (43.9)	17.4 (45.5)	18.4 (55.2)
3ORF1	24.8 (50.4)	23.9 (46.9)	26.5 (47.8)	26.0 (48.8)	29.3 (56.9)	25.0 (53.3)	22.0 (48.0)	26.0 (50.4)	19.5 (49.6)	17.7 (52.9)	20.7 (44.8)	23.6 (48.0)	28.0 (55.9)	20.9 (47.0)	17.4 (38.8)	17.2 (39.3)	15.0 (36.7)	16.1 (49.2)
5ORF1	14.1 (48.4)	19.4 (47.6)	16.1 (47.6)	23.1 (49.6)	23.0 (51.6)	27.2 (52.8)	19.5 (49.6)	23.0 (52.4)	17.3 (49.6)	17.6 (52.0)	16.8 (38.7)	23.0 (46.0)	22.8 (49.6)	15.0 (46.7)	19.7 (44.3)	14.8 (47.5)	16.5 (45.7)	13.7 (51.6)
9ORF1	9.6 (41.6)	16.0 (41.6)	17.6 (39.5)	21.8 (46.8)	29.8 (57.3)	26.7 (53.3)	19.5 (49.6)	21.0 (50.0)	19.4 (48.4)	18.9 (53.3)	14.1 (43.0)	21.0 (45.2)	25.8 (55.7)	14.2 (48.3)	12.1 (40.3)	20.8 (45.0)	14.8 (37.7)	17.2 (51.6)
PAV-3 E4 ORF3[b]	20.2 (52.4)	22.6 (51.6)	21.8 (46.8)	24.8 (47.9)	27.2 (52.0)	24.0 (55.2)	35.0 (58.5)	30.6 (55.6)	28.2 (57.3)	31.2 (53.6)	17.5 (41.7)	32.8 (52.0)	28.8 (55.2)	17.6 (49.6)	19.7 (41.0)	20.9 (51.3)	17.4 (44.6)	20.7 (50.4)

[a] Determined for full-length E4 ORF1 and truncated dUTPase core sequences (see Materials and Methods) as described in the footnote to Table 1. Abbreviations not given in the text: *H. influenzae*, *Haemophilus influenzae*; *C. burnetii*, *Coxiella burnetii*; SRV-1, simian retrovirus type 1; MMTV, mouse mammary tumor virus; FIV, feline immunodeficiency virus; HSV-1, herpes simplex virus type 1. The boxes represent E4 ORF1:dUTPase pairs having some of the highest sequence similarities for these analyses (see text).

[b] Included because it may be an E4 ORF1 homolog (see text).

tectable dUTPase activity for the recombinant E4 ORF1 proteins in the enzyme assays was further supported by the fact that, in contrast to mammalian cells overexpressing *Escherichia coli* dUTPase (12), transformed CREF cells overexpressing the 9ORF1 protein did not show elevated or altered levels of dUTPase activity (data not shown). Together, these findings strongly suggested that E4 ORF1 transforming proteins are not dUTPase enzymes.

Adenovirus E4 ORF1 proteins may be evolutionarily related to dUTPase enzymes. As mentioned above, an avian adenovirus (CELO) encodes a putative dUTPase which, compared to 17 other dUTPases, was among four having the highest sequence similarity with the E4 ORF1 proteins (Table 2). Significantly, the gene for this putative dUTPase has been proposed to be situated in a genomic location analogous to that of the human adenovirus E4 ORF1 genes (1, 16). While dUTPase enzymatic activity has not been demonstrated for the CELO protein, it displayed extensive sequence similarity with the human dUTPase protein (Table 3) and preserved all of the dUTPase conserved motifs (Fig. 5). Therefore, we tested a GST-CELO protein construct for dUTPase enzymatic activity. As opposed to the GST-E4 ORF1 proteins, the GST-CELO protein possessed significant dUTPase enzymatic activity, behaving like GST-human dUTPase in the enzyme assays (Fig. 6A). Additionally, as reported for other dUTPases (48, 87, 89), the enzymatic activity of the CELO protein was competed almost completely by unlabeled dUTP, but not UTP or dATP, further indicating that dUTP was a specific substrate for this enzyme (Fig. 6B). From these results, we concluded that the CELO protein was a genuine dUTPase enzyme. More important, considering the sequence similarity between E4 ORF1 and dUTPase proteins, this finding hinted to the possibility that human adenovirus E4 ORF1 genes evolved from an ancestral adenovirus dUTPase gene.

Adenovirus E4 ORF1 proteins may have structural similarities with dUTPase enzymes. Sequence similarity between proteins may indicate structural and/or functional relatedness. As described above, despite having sequence similarity with dUTPases, the E4 ORF1 proteins lacked detectable dUTPase enzymatic activity. We also failed to detect a physical association between the 9ORF1 and human dUTPase proteins; in addition, we did not observe cellular growth-transforming activity for expression plasmids encoding human, avian adenovirus CELO, vaccinia virus, or mouse mammary tumor virus dUTPase genes (data not shown). Therefore, utilizing computer programs capable of predicting protein secondary structures with reasonable accuracy, we next addressed the possibility that the E4 ORF1 proteins are structurally related to the dUTPase proteins. The rationale for these analyses was that similar types and arrangements of protein secondary structures for polypeptides might indicate related tertiary structures.

The crystal structure of *E. coli* dUTPase has been solved and shows an interdigitating trimer, with each subunit adopting an eight-stranded jelly roll conformation (13, 52). Using this detailed structural information, we first evaluated the accuracy of three different computer programs (73, 79, 92) to correctly predict the known secondary structures of this relevant dUTPase. While all programs identified most of the known secondary structures, the PHD program (73) predicted secondary structures with highest accuracy in the test comparisons with the *E. coli* dUTPase protein. The *E. coli* dUTPase protein consists of eight major β-strands (β1 to β8), of which β5 is separated into two smaller β-strands by a short loop region. Also present in the molecule are two shorter β-strands (β2/3 and β6/7) and an α-helical region (α4/5) between β4 and β5. In the test analysis, the PHD algorithm (i) properly identified β1,

TABLE 3. Sequence similarities among 18 different dUTPase proteins

dUTPase protein from:	% Amino acid identity[a] (% amino acid similarity) to dUTPase protein from:																	
	Prokaryote				Eukaryote				DNA virus					Retrovirus				
	E. coli	*H. influenzae*	*C. burnetii*	*M. leprae*	*C. albicans*	*S. cerevisiae*	*H. sapiens*	*L. esculentum*	Vaccinia virus	Shope fibroma virus	HSV-1	Orf virus	Avian adenovirus CELO	Visna virus	SRV-1	MMTV	EIAV	FIV
HSV-1	33.1 (58.3)	35.5 (55.8)	38.7 (60.6)	35.9 (55.6)	33.6 (55.7)	30.3 (53.5)	36.3 (59.3)	37.9 (52.9)	33.6 (53.3)	34.3 (54.5)	**100**	40.3 (61.2)	29.1 (51.8)	29.0 (58.8)	33.1 (55.9)	39.3 (59.3)	29.7 (54.3)	30.4 (54.8)
FIV	30.6 (59.0)	29.1 (56.7)	34.6 (54.9)	32.1 (58.6)	29.0 (55.2)	27.9 (59.3)	32.9 (56.4)	32.4 (61.3)	35.0 (60.8)	32.4 (59.7)		30.7 (60.0)	28.3 (59.3)	44.9 (68.4)	35.0 (56.2)	37.1 (55.7)	45.8 (71.8)	**100**
EIAV	24.3 (60.0)	27.9 (58.6)	29.5 (56.1)	25.4 (54.9)	28.8 (52.1)	28.3 (51.7)	32.8 (62.8)	28.3 (60.1)	31.5 (53.8)	24.5 (55.4)		28.8 (55.5)	26.0 (54.1)	42.6 (66.9)	30.4 (58.5)	29.2 (55.5)	**100**	
MMTV	35.8 (61.2)	34.3 (59.1)	38.8 (60.4)	31.9 (56.5)	37.6 (54.9)	31.4 (52.9)	36.6 (56.7)	34.6 (54.9)	32.6 (53.3)	33.3 (51.9)		37.8 (55.6)	37.0 (57.0)	32.1 (56.5)	46.5 (59.2)	**100**		
SRV-1	29.2 (55.5)	31.2 (50.4)	35.3 (53.7)	37.8 (57.4)	34.3 (56.9)	33.8 (58.1)	32.6 (53.3)	32.4 (55.1)	32.8 (52.6)	34.1 (54.1)		33.1 (53.7)	29.1 (49.6)	30.5 (55.7)	**100**			
Visna virus	29.9 (61.9)	34.6 (63.2)	33.1 (55.6)	28.7 (57.4)	32.4 (59.6)	30.4 (54.8)	30.1 (55.1)	34.1 (60.7)	32.4 (57.4)	26.7 (50.4)		27.9 (52.2)	31.6 (59.6)	**100**				
Avian adenovirus CELO	32.0 (59.2)	33.6 (56.8)	33.6 (55.5)	40.8 (62.6)	51.0 (69.7)	52.8 (70.4)	59.6 (78.0)	57.2 (76.6)	50.0 (68.5)	50.7 (70.7)		58.9 (73.3)	**100**					
Orf virus	32.9 (54.1)	34.9 (54.1)	37.5 (54.9)	38.1 (61.2)	53.7 (68.7)	49.3 (65.1)	69.5 (78.7)	56.8 (70.5)	53.4 (71.2)	51.4 (72.5)		**100**						
Shope fibroma virus	37.3 (58.5)	38.0 (61.3)	33.8 (53.5)	39.4 (59.2)	44.2 (70.3)	41.5 (63.4)	52.1 (72.9)	47.2 (70.4)	57.4 (75.2)	**100**								
Vaccinia virus	31.5 (55.5)	37.2 (56.6)	38.3 (53.9)	40.4 (58.2)	45.9 (68.5)	42.8 (64.8)	63.6 (79.3)	52.8 (70.8)	**100**									
L. esculentum	35.9 (60.0)	37.2 (61.4)	35.9 (58.6)	41.8 (59.6)	56.2 (73.3)	50.0 (67.1)	61.4 (75.0)	**100**										
H. sapiens	32.6 (58.9)	37.6 (58.9)	40.0 (58.6)	39.7 (63.1)	60.3 (72.3)	54.2 (70.7)	**100**											
S. cerevisiae	30.3 (61.4)	32.4 (60.0)	37.2 (60.0)	34.2 (61.6)	63.3 (77.6)	**100**												
C. albicans	36.1 (63.2)	42.3 (63.4)	38.2 (59.0)	38.3 (61.7)	**100**													
M. leprae	34.9 (61.1)	35.8 (58.8)	37.8 (58.8)	**100**														
C. burnetii	61.6 (81.5)	68.7 (82.0)	**100**															
H. influenzae	75.5 (88.7)	**100**																
E. coli	**100**																	

[a] Determined for dUTPase core sequences (see Materials and Methods) as described in the footnote to Table 1. See Table 2, footnote a, for abbreviations. The boxes represent dUTPase:dUTPase pairs having some of the lowest sequence similarities for these analyses (see text).

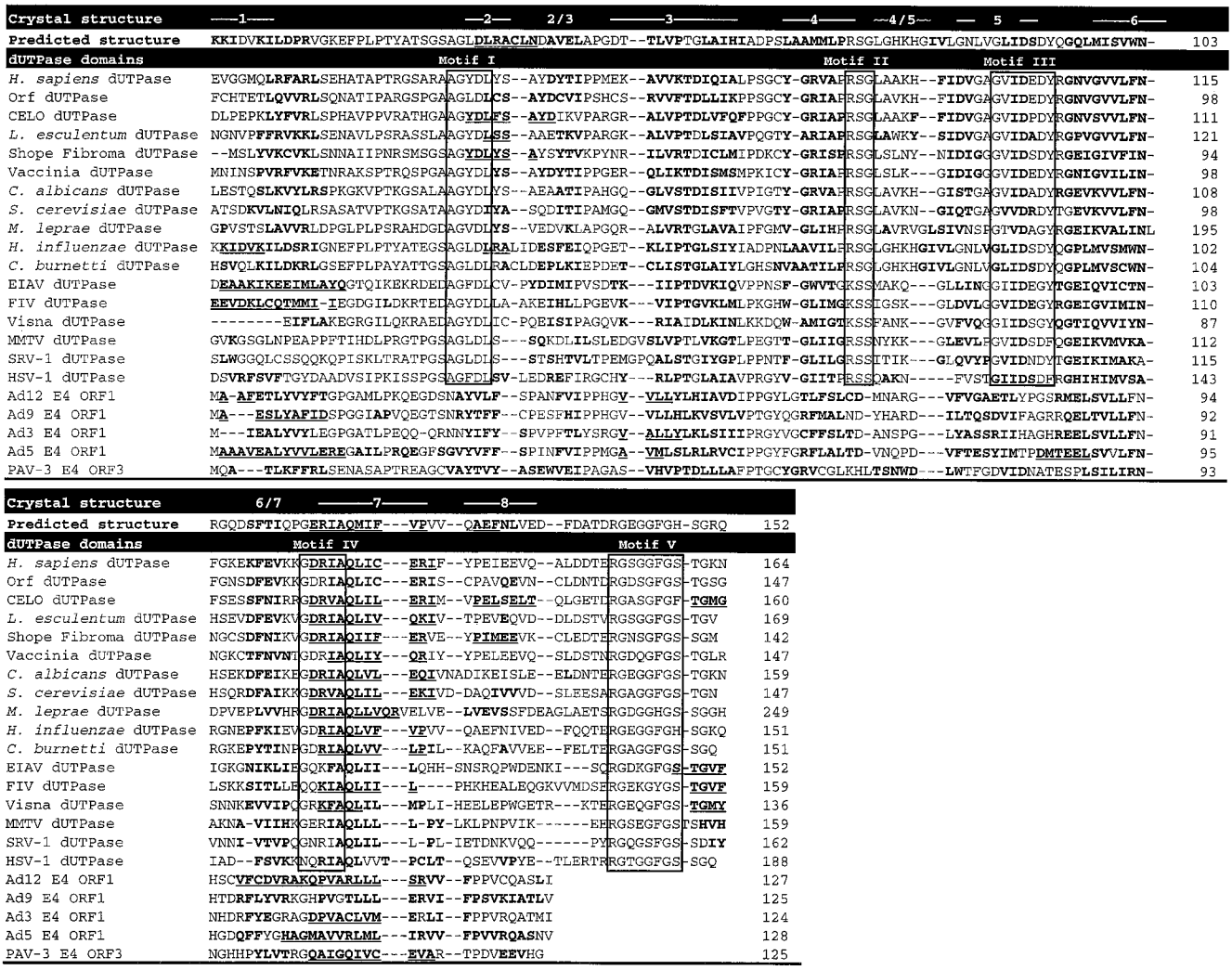


FIG. 5. Multiple alignment of adenovirus E4 ORF1 and 18 different dUTPase sequences showing (i) secondary structures (β-strand, boldface amino acid residues; α-helix, boldface and underlined amino acid residues) predicted by the PHD secondary structure prediction program (73) and (ii) the locations of the conserved dUTPase motifs I to V (dUTPase domains; boxed regions) (51, 57). The accuracy of the PHD program is illustrated in the upper two rows, which compare the types and locations of known secondary structures (β-strand, straight line; α-helix, wavy line) from the crystal structure of *E. coli* dUTPase (crystal structure) (13, 52) with those predicted by the PHD program (predicted structure). The amino acid sequences of human adenovirus E4 ORF1 proteins were aligned with dUTPase core sequences (see Materials and Methods) using the PIMA multiple sequence alignment algorithm (77, 78). A putative E4 ORF1 homologue, PAV-3 E4 ORF3 (see Results), is also included in these analyses. See Table 2, footnote *a*, for abbreviations.

β2/3, β3, β4, β5, β6, and β6/7, (ii) properly located β2, β7, and β8 but improperly designated them as completely or partially α-helical, and (iii) failed to detect α4/5 (Fig. 5). Thus, while incorrectly labeling some β-strand regions as α-helical, this particular program located all but one known secondary structure of the *E. coli* dUTPase protein.

Prompted by these results, we decided to use the PHD program to predict the positions of secondary structures within the E4 ORF1 proteins and dUTPase core sequences of the remaining 17 dUTPase proteins. Despite wide variations in amino acid sequence similarity (Table 3), all dUTPases produced predicted secondary structure arrangements strikingly similar to that of *E. coli* dUTPase (Fig. 5), suggesting that these enzymes will be found to have common structures. More important, the predicted secondary structure organizations of the human adenovirus E4 ORF1 proteins were also highly concordant with those of the dUTPase proteins. While the predicted secondary structures of the E4 ORF1 proteins deviated slightly from those of dUTPases by showing continuous β5 strands rather than segmented ones like that of *E. coli* dUTPase and, in some cases, α-helical structure within β3, β6, and β6/7 of *E. coli* dUTPase, in general, the predicted E4 ORF1 secondary structures matched well with those of the dUTPases in both type and location (Fig. 5). The carboxy-terminal residues of *E. coli* dUTPase, corresponding precisely to the residues missing from E4 ORF1 proteins (Fig. 5), are not visible in the crystal structure and are presumed to be disordered (13, 52). This fact suggests that the carboxy-terminal region of dUTPases may not be an integral component of their core structures. Therefore, although lacking the carboxy-terminal region of dUTPases, the E4 ORF1 proteins could still theoretically adopt a tertiary structure similar to that of *E. coli* dUTPase.

A putative E4 ORF1 homolog is encoded within the E4 region of PAV-3. During the course of these studies, we also found that a predicted PAV-3 E4 region-encoded protein (E4 ORF3), having no known function, showed sequence (Tables 1

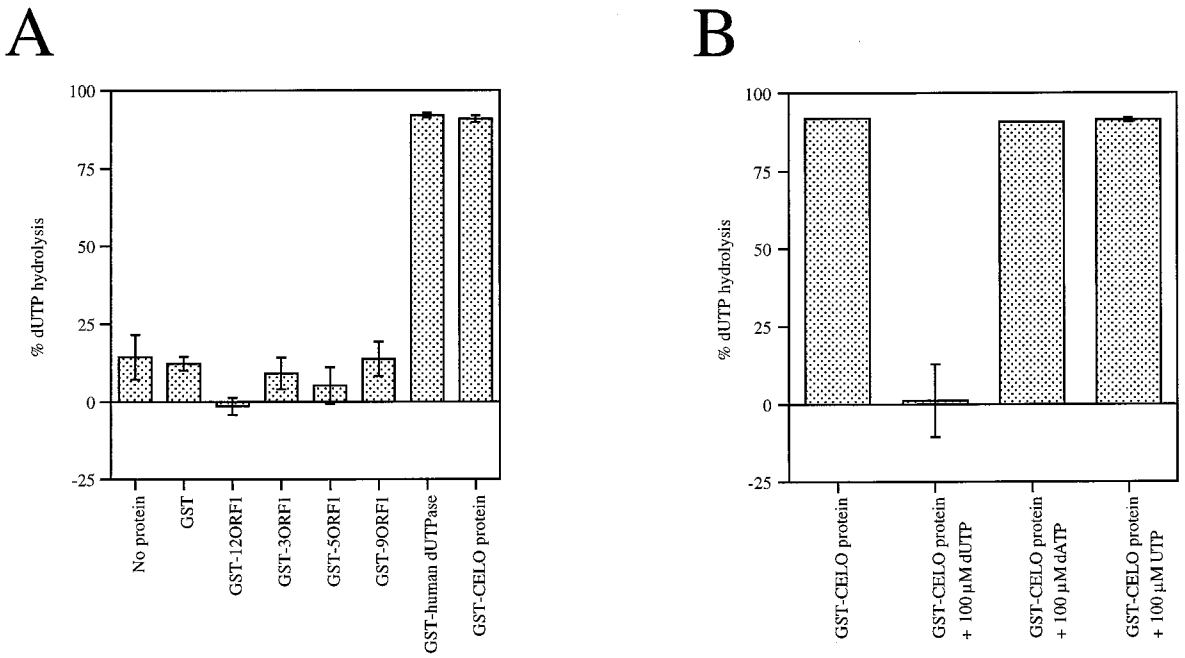


FIG. 6. (A) dUTPase enzyme assays with 50 ng of GST constructs containing each adenovirus E4 ORF1, human dUTPase, or the putative dUTPase of avian adenovirus CELO (CELO protein). (B) dUTPase enzyme assays with 15 ng of GST-CELO protein in the presence or absence of 100 μM cold nucleotide (dUTP, UTP, or dATP) competitor. Enzyme assays measured the amount of deoxy[5-^{3}H]uridine 5′-triphosphate (dUTP) that remained unhydrolyzed after incubation for 5 min at 37°C with the indicated protein. Percent dUTP hydrolysis values for samples were calculated as described in Materials and Methods. The means of three samples, with error bars representing the standard deviations, are shown.

and 2) and predicted structural similarities (Fig. 5) with both adenovirus E4 ORF1 and dUTPase proteins. Interestingly, while exhibiting higher sequence similarity to eukaryotic, avian adenovirus CELO, and poxvirus dUTPases than to the human adenovirus E4 ORF1 proteins, the predicted PAV-3 E4 ORF3 polypeptide exhibited the physical characteristics of an E4 ORF1 protein by its length and lack of preservation of conserved dUTPase motifs (Fig. 5). Closer relatedness to the E4 ORF1 proteins was further implied by phylogenetic analyses, which placed the PAV-3 E4 ORF3 and E4 ORF1 proteins together within one branch while logically separating the dUTPase sequences of prokaryotes, retroviruses, and eukaryotes/DNA viruses into three other distinct branches (Fig. 7). While it is not known whether PAV-3 E4 ORF3 possesses transforming activity, these findings may indicate that it is an E4 ORF1 protein homolog having higher sequence similarity to dUTPase enzymes than to human adenovirus E4 ORF1 proteins.

DISCUSSION

One goal of this study was to determine whether all of the different human adenovirus E4 ORF1 proteins had transforming activities in cells. In a cell line capable of expressing high levels of most of these proteins, we demonstrated that adenovirus E4 ORF1s from subgroups A to D do indeed possess similar growth-transforming potentials, suggesting that these related proteins carry out a common function in the life cycle of human adenoviruses. The identification of a putative E4 ORF1 homolog encoded by a porcine adenovirus may further indicate wider importance of this E4 function for adenoviruses from other species as well. Additionally, the fact that Ad5 E4 ORF1 was found to have potent cellular growth-transforming activity is relevant to the design of replication-defective Ad5 vectors, as their use for gene therapy protocols in humans has raised numerous safety concerns, including the possibility of oncogenic transformation (49). While E1 region oncogenes are deleted from adenovirus vectors, the results presented here for E4 ORF1 and for E4 ORF6 by others (20) strongly argue that routine removal of E4 region sequences from these vectors would be prudent (9, 91).

The work described herein also suggests that the lack of transforming activity observed for non-subgroup D virus E4 ORF1s in CREF cells (45) may be due to inadequate expression of these viral genes. The expression deficiencies were not specific to CREF cells, as similar results were obtained in the C127 murine fibroblast cell line. These observations contrasted findings with several human cell lines, in which expression of most non-subgroup D virus E4 ORF1s was comparable to that of 9ORF1. The differences between the rodent and human cell lines may denote a specific block in the expression of non-subgroup D virus E4 ORF1 genes in some species or cell types. While the basis for the expression deficiencies of these E4 ORF1 genes was not determined, such defects could explain the difference in oncogenic tropism between Ad9 and other oncogenic adenoviruses (2, 3, 42, 85). The facts that Ad9 requires E4 ORF1 for its oncogenicity (45) and non-subgroup D adenoviruses do not (28) may indicate that E4 ORF1 actually determines the tropism of Ad9 tumorigenicity for the mammary gland. If this were correct, deficient expression of E4 ORF1 proteins in rodent cells might, at least in part, account for non-subgroup D adenoviruses failing to elicit mammary tumors. Conversely, this scenario would also imply that the Ad9 E1 region is nononcogenic, as Ad9 does not elicit the nonmammary tumors observed for non-subgroup D adenoviruses. In support of the latter prediction, Ad9 E1 region-transformed CREF cells do not show increased oncogenicity in animals (88).

An additional goal of this study was to determine whether

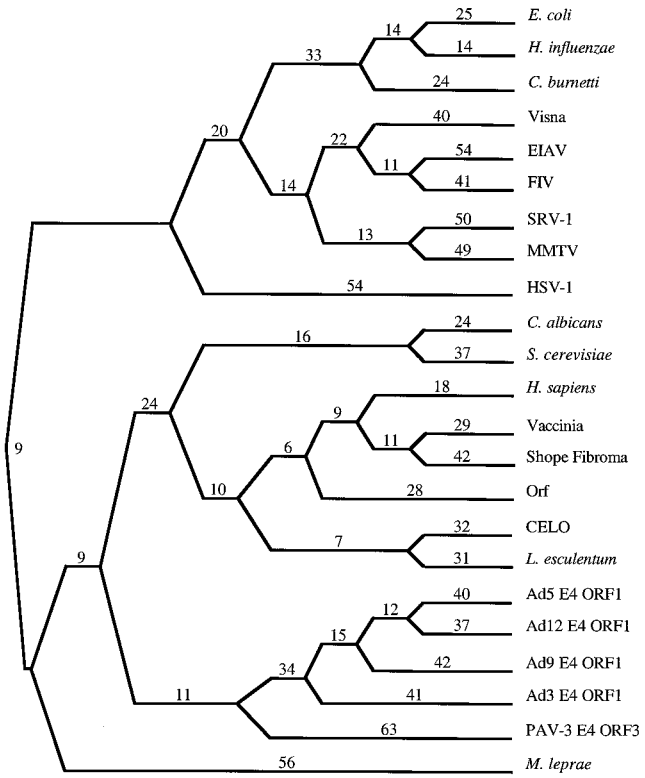


FIG. 7. Phylogenetic tree of adenovirus E4 ORF1 and 18 different dUTPase proteins. Phylogenetic analyses were performed by the parsimony method using the PAUP software program (83) with full-length E4 ORF1 sequences and truncated dUTPase core sequences (see Materials and Methods). The numbers refer to branch lengths which are proportional to the estimated number of amino acid residue substitutions. The heuristic search procedure with the midpoint rooting method was used to generate the tree with the following parameters: the maximum number of trees held in memory was set to 100, branches of length 0 were collapsed, stepwise addition of protein sequences to the tree was by the "simple" algorithm, and branch swapping was performed on each minimal-length tree by the tree bisection-reconnection option. See Table 2, footnote *a*, for abbreviations.

the E4 ORF1 family of viral transforming proteins was related to a known protein. With regard to this possibility, we found limited sequence and predicted structural similarities between human adenovirus E4 ORF1 proteins and a wide variety of dUTPase enzymes. While further experiments are necessary to determine whether these proteins actually have similar structures, numerous examples are known for proteins having minimal amino acid sequence similarity yet possessing highly related three-dimensional structures (37, 38, 54, 70). For many viral proteins, however, sequence similarity with a cellular protein additionally reflects retention of a functional trait. This is best illustrated by some poxvirus and herpesvirus proteins which show a high degree of both sequence and functional similarity with cellular enzymes and immune modulatory factors (11, 72). More relevant to this study, the adenovirus E1B 19K oncoprotein has been shown to exhibit sequence similarity with the cellular proto-oncogene product, Bcl-2, and has indistinguishable anti-apoptotic function (17). In some cases, however, only a small conserved domain of a cellular protein may be recognizable within viral oncoproteins. Most notably, the LXCXE motif, which mediates binding of some cellular proteins to the pRb family of tumor suppressors, is also found in the adenovirus E1A, simian virus 40 large T antigen, and papillomavirus E7 oncoproteins (14, 19, 46, 55, 90). Therefore, even though the sequence similarity reported here between the adenovirus E4 ORF1 transforming proteins and dUTPase enzymes did not lead to identification of a common function, it is possible that further studies will reveal an undiscovered activity shared by these two related proteins.

It must also be considered that E4 ORF1 and dUTPase proteins are truly unrelated functionally, although this would not preclude these proteins from being evolutionarily linked. Avian adenovirus CELO has been proposed to encode a dUTPase (16); in support of this claim, results described in this report formally demonstrate that the putative avian adenovirus CELO dUTPase is an authentic dUTPase enzyme. Therefore, an evolutionary relationship between E4 ORF1 and dUTPase proteins may be implied by the fact that the avian adenovirus CELO dUTPase is encoded within a genomic region analogous to that of human adenovirus E4 ORF1 (16). Given that E4 ORF1 and dUTPase proteins may not share a common function, precedence for the existence of evolutionarily related proteins that are structurally similar yet functionally distinct comes from studies of c-type lysozymes and α-lactalbumins (7, 60, 69, 82). Perhaps the human adenovirus E4 ORF1s have similarly evolved divergently from an ancestral dUTPase gene, with E4 ORF1 proteins retaining dUTPase structural features and acquiring novel transforming activity. In this respect, evolutionary precursors of avian adenovirus CELO dUTPase and PAV-3 E4 ORF3 may represent transitional intermediates between the initial transduction of a cellular dUTPase gene into an adenovirus genome and its subsequent evolution into the E4 ORF1 transforming genes. Such a model would argue against the possible convergent evolution of the adenovirus E4 ORF1 and dUTPase proteins.

As is sometimes the case for related viruses, avian adenovirus CELO and human adenoviruses differ with respect to encoding dUTPase enzymes (10, 22, 57, 75). Results from several systems suggest that having a dUTPase gene significantly influences viral tropism. Studies of dUTPase mutants of EIAV, feline immunodeficiency virus, caprine arthritis-encephalitis virus, and herpes simplex virus type 1 indicate that viral dUTPases are dispensable in cells having high endogenous cellular dUTPase levels but are required for efficient viral replication in cells with low dUTPase levels, such as terminally differentiated or nondividing cells (25, 68, 84, 86, 87). Consequently, the presence of a dUTPase gene in avian adenovirus CELO, but not human adenoviruses, may indicate differences in the tropism and pathogenesis of these viruses or, perhaps, that human adenoviruses utilize a novel mechanism to circumvent the detrimental consequences of dUTP accumulation during DNA replication (53, 71, 80).

Although no similar function was found for the E4 ORF1 and dUTPase proteins, the results of this study suggested that these related proteins may share a common structure. The three-dimensional structure of the *E. coli* dUTPase protein offers no obvious advantages for a viral transforming protein, but it is noteworthy that this enzyme, like most dUTPases, exists as a homotrimer (4, 5, 13, 18, 35). Because many viral oncoproteins are known to homo-oligomerize, this particular structural feature of dUTPase polypeptides could potentially benefit the E4 ORF1 transforming protein. With regard to the functional significance of homo-oligomerization by viral oncoproteins, perhaps the best-understood example is that of the bovine papillomavirus E5 protein, for which transforming activity relies on protein dimerization for induction of platelet-derived growth factor receptor subunit coupling, *trans* autophosphorylation, and subsequent ligand-independent signaling (61, 64, 67). The simian virus 40 large T antigen, human papillomavirus E7, and human adenovirus E1B 55K oncoproteins

may also form higher-order homo-oligomeric structures (6, 29, 59, 63), although the precise role of these complexes in transformation is not as well understood. We are presently examining interactions of the E4 ORF1 proteins with cellular proteins; in light of the predicted structural similarity with dUTPase proteins, future experiments will also be directed at determining whether the E4 ORF1 proteins homo-oligomerize. The examination of hetero- and homo-oligomerization by the E4 ORF1 transforming proteins in cells may lead to a detailed model for the molecular mechanisms of these novel transforming proteins.

ACKNOWLEDGMENTS

We thank Jeff Rosen for helpful discussions, Darby Thomas for comments on the manuscript, and Caroline Petrin for assistance with figures. We also acknowledge the following individuals for providing reagents: Olivia Pereira-Smith and Jim Smith for TE85 cells, Joseph Weber for avian adenovirus CELO, Jerry Shaack for Ad3, and Evan M. McIntosh for plasmid pHUMDUT1.

R.S.W. was the recipient of a National Science Foundation graduate research fellowship. This work was supported by grant CA 58541 from the National Cancer Institute.

REFERENCES

1. **Akopian, T. A., E. N. Kaverina, B. S. Naroditsky, and T. I. Tikhonenko.** 1992. Nucleotide sequence analysis of the avian adenovirus CELO (FAV-1) DNA fragment (92–100%). Mol. Gen. Mikrobiol. Virol. **11:**19–23.
2. **Ankerst, J., and N. Jonsson.** 1989. Adenovirus type 9-induced tumorigenesis in the rat mammary gland related to sex hormonal state. J. Natl. Cancer Inst. **81:**294–298.
3. **Ankerst, J., N. Jonsson, L. Kjellen, E. Norrby, and H. O. Sjogren.** 1974. Induction of mammary fibroadenomas in rats by adenovirus type 9. Int. J. Cancer **13:**286–290.
4. **Bergman, A. C., O. Bjornberg, J. Nord, P. O. Nyman, and A. M. Rosengren.** 1994. The protein p30, encoded at the gag-pro junction of mouse mammary tumor virus, is a dUTPase fused with a nucleocapsid protein. Virology **204:**420–424.
5. **Bergman, A. C., O. Bjornberg, J. Nord, A. M. Rosengren, and P. O. Nyman.** 1995. dUTPase from the retrovirus equine infectious anemia virus: high-level expression in *Escherichia coli* and purification. Protein Expression Purif. **6:**379–387.
6. **Bradley, M. K., J. D. Griffin, and D. M. Livingston.** 1982. Relationship of oligomerization to enzymatic and DNA-binding properties of the SV40 large T antigen. Cell **28:**125–134.
7. **Brew, K., T. C. Vanaman, and R. L. Hill.** 1967. Comparison of the amino acid sequence of bovine α-lactalbumin and hens egg white lysozyme. J. Biol. Chem. **242:**3747–3749.
8. **Bridge, E., and G. Ketner.** 1989. Redundant control of adenovirus late gene expression by early region 4. J. Virol. **63:**631–638.
9. **Brough, D. E., A. Lizonova, C. Hsu, V. A. Kulesa, and I. Kovesdi.** 1996. A gene transfer vector-cell line system for complete functional complementation of adenovirus early regions E1 and E4. J. Virol. **70:**6497–6501.
10. **Broyles, S. S.** 1993. Vaccinia virus encodes a functional dUTPase. Virology **195:**863–865.
11. **Buller, R. M. L., and G. J. Palumbo.** 1991. Poxvirus pathogenesis. Microbiol. Rev. **55:**80–122.
12. **Canman, C. E., E. H. Radany, L. A. Parsels, M. A. Davis, T. S. Lawrence, and J. Maybaum.** 1994. Induction of resistance to fluorodeoxyuridine cytotoxicity and DNA damage in human tumor cells by expression of *Escherichia coli* deoxyuridinetriphosphatase. Cancer Res. **54:**2296–2298.
13. **Cedergren-Zeppezauer, E. S., G. Larsson, P. O. Nyman, Z. Dauter, and K. S. Wilson.** 1992. Crystal structure of a dUTPase. Nature **355:**740–743.
14. **Chellappan, S., V. B. Kraus, B. Kroger, K. Munger, P. M. Howley, W. C. Phelps, and J. R. Nevins.** 1992. Adenovirus E1A, simian virus 40 tumor antigen, and human papillomavirus E7 proteins share the capacity to disrupt the interaction between transcription factor E2F and the retinoblastoma gene product. Proc. Natl. Acad. Sci. USA **89:**4549–4553.
15. **Chen, M., and M. S. Horwitz.** 1990. Replication of an adenovirus type 34 mutant DNA containing tandem reiterations of the inverted terminal repeat. Virology **179:**567–575.
16. **Chiocca, S., R. Kurzbauer, G. Schaffner, A. Baker, V. Mautner, and M. Cotten.** 1996. The complete DNA sequence and genomic organization of the avian adenovirus CELO. J. Virol. **70:**2939–2949.
17. **Chiou, S. K., C. C. Tseng, L. Rao, and E. White.** 1994. Functional complementation of the adenovirus E1B 19-kilodalton protein with Bcl-2 in the inhibition of apoptosis in infected cells. J. Virol. **68:**6553–6566.
18. **Climie, S., T. Lutz, J. Rdul, M. Sumner-Smith, E. Vandenberg, and E. McIntosh.** 1994. Expression of trimeric human dUTP pyrophosphatase in *Escherichia coli* and purification of the enzyme. Protein Expression Purif. **5:**252–258.
19. **Defeo-Jones, D., P. S. Huang, R. E. Jones, K. M. Haskell, G. A. Vuocolo, M. G. Hanobik, H. E. Huber, and A. Oliff.** 1991. Cloning of cDNAs for cellular proteins that bind to the retinoblastoma gene product. Nature **352:** 251–254.
20. **Dobner, T., N. Horikoshi, S. Rubenwolf, and T. Shenk.** 1996. Blockage by adenovirus E4orf6 of transcriptional activation by the p53 tumor suppressor. Science **272:**1470–1473.
21. **Duker, N. J., and C. L. Grant.** 1980. Alterations in the levels of deoxyuridine triphosphatase, uracil-DNA glycosylase and AP endonuclease during the cell cycle. Exp. Cell Res. **125:**493–497.
22. **Elder, J. H., D. L. Lerner, L. C. Hasselkus, D. J. Fontenot, E. Hunter, P. A. Luciw, R. C. Montelaro, and T. R. Phillips.** 1992. Distinct subsets of retroviruses encode dUTPase. J. Virol. **66:**1791–1794.
23. **El Hajj, H., H. Zhang, and B. Weiss.** 1988. Lethality of a *dut* (deoxyuridine triphosphatase) mutation in *Escherichia coli*. J. Bacteriol. **170:**1069–1075.
24. **Field, J., J. I. Nikawa, D. Broek, B. MacDonald, L. Rodgers, I. A. Wilson, R. A. Lerner, and M. Wigler.** 1988. Purification of a *RAS*-responsive adenylyl cyclase complex from *Saccharomyces cerevisiae* by use of an epitope addition method. Mol. Cell. Biol. **8:**2159–2165.
25. **Fisher, F. B., and V. G. Preston.** 1986. Isolation and characterisation of herpes simplex virus type 1 mutants which fail to induce dUTPase activity. Virology **148:**190–197.
26. **Fisher, P. B., L. E. Babiss, I. B. Weinstein, and H. S. Ginsberg.** 1982. Analysis of type 5 adenovirus transformation with a cloned rat embryo cell line (CREF). Proc. Natl. Acad. Sci. USA **79:**3527–3531.
27. **Gadsden, M. H., E. M. McIntosh, J. C. Game, P. J. Wilson, and R. H. Haynes.** 1993. dUTP pyrophosphatase is an essential enzyme in Saccharomyces cerevisiae. EMBO J. **12:**4425–4431.
28. **Graham, F. L.** 1984. Transformation by and oncogenicity of human adenoviruses, p. 339–398. *In* H. S. Ginsberg (ed.), The adenoviruses. Plenum Press, New York, N.Y.
29. **Grand, R. J. A., T. Mustoe, S. Roberts, and P. H. Gallimore.** 1995. The quartenary structure of the adenovirus 12 early region 1B 54K protein. Virology **207:**255–259.
30. **Halbert, D. N., J. R. Cutt, and T. Shenk.** 1985. Adenovirus early region 4 encodes functions required for efficient DNA replication, late gene expression, and host cell shutoff. J. Virol. **56:**250–257.
31. **Herisse, J., M. Rigolet, S. D. de Dinechin, and F. Galibert.** 1981. Nucleotide sequence of adenovirus 2 DNA fragment encoding for the carboxylic region of fiber protein and the entire E4 region. Nucleic Acids Res. **9:**4023–4042.
32. **Hinds, P. W., C. A. Finlay, R. S. Quartin, S. J. Baker, E. R. Fearon, B. Vogelstein, and A. J. Levine.** 1990. Mutant p53 cDNAs from human colorectal carcinomas can cooperate with ras in transformation of rat cells. Cell Growth Differ. **1:**571–580.
33. **Hogenkamp, T., and H. Esche.** 1990. Nucleotide sequence of the right 10% of adenovirus type 12 DNA encoding the entire region E4. Nucleic Acids Res. **18:**3065–3066.
34. **Hokari, S., M. Hasegawa, Y. Sakagishi, and G. Kikuchi.** 1987. Deoxyuridine triphosphate nucleotidohydrolase activity and its correlation with multiplication of erythroid cells in rat spleen. Biochem. Int. **14:**851–857.
35. **Hokari, S., and Y. Sakagishi.** 1987. Purification and characterization of deoxyuridine triphosphate nucleotidohydrolase from anemic rat spleen: a trimer composition of the enzyme protein. Arch. Biochem. Biophys. **253:** 350–356.
36. **Hokari, S., Y. Sakagishi, and K. Tsukada.** 1982. Enhanced activity of deoxyuridine 5′-triphosphatase in regenerating rat liver. Biochem. Biophys. Res. Commun. **108:**95–101.
37. **Holm, L., and C. Sander.** 1994. Searching protein structure databases has come of age. Proteins **19:**165–173.
38. **Holm, L., and C. Sander.** 1995. Evolutionary link between glycogen phosphorylase and a DNA modifying enzyme. EMBO J. **14:**1287–1293.
39. **Horwitz, M. S.** 1990. Adenoviridae and their replication, p. 1679–1721. *In* B. N. Fields, D. M. Knipe, et al. (ed.), Virology, 2nd ed. vol. 2. Raven Press, Ltd., New York, N.Y.
40. **Horwitz, M. S.** 1990. Adenoviruses, p. 1723–1740. *In* B. N. Fields, D. M. Knipe, et al. (ed.), Virology, 2nd ed., vol. 2. Raven Press, Ltd., New York, N.Y.
41. **Huang, M. M., and P. Hearing.** 1989. Adenovirus early region 4 encodes two gene products with redundant effects in lytic infection. J. Virol. **63:**2605–2615.
42. **Javier, R., K. Raska, Jr., G. J. Macdonald, and T. Shenk.** 1991. Human adenovirus type 9-induced rat mammary tumors. J. Virol. **65:**3192–3202.
43. **Javier, R., K. Raska, Jr., and T. Shenk.** 1992. Requirement for the adenovirus type 9 E4 region in production of mammary tumors. Science **257:**1267–1271.
44. **Javier, R., and T. Shenk.** 1996. Mammary tumors induced by human adenovirus type 9: a role for the early region 4 gene. Breast Cancer Res. Treat. **39:**57–67.
45. **Javier, R. T.** 1994. Adenovirus type 9 E4 open reading frame 1 encodes a

transforming protein required for the production of mammary tumors in rats. J. Virol. **68:**3917–3924.
46. **Kaelin, W. G. J., W. Krek, W. R. Sellers, J. A. DeCaprio, F. Ajchenbaum, C. S. Fuchs, T. Chittenden, Y. Li, P. J. Farnham, M. A. Blanar, D. M. Livingston, and E. K. Flemington.** 1992. Expression cloning of a cDNA encoding a retinoblastoma-binding protein with E2F-like properties. Cell **70:**351–364.
47. **Kingston, R. E., C. A. Chen, and H. Okayama.** 1990. Calcium phosphate transfection, p. 9.1.1–9.1.9. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology. Greene Publishing Associates and Wiley-Interscience, New York, N.Y.
48. **Koppe, B., A. L. Menendez, and S. Oroszlan.** 1994. Expression and purification of the mouse mammary tumor virus gag-pro transframe protein p30 and characterization of its dUTPase activity. J. Virol. **68:**2313–2319.
49. **Kurth, R.** 1995. Risk potential of the chromosomal insertion of foreign DNA. Ann. N. Y. Acad. Sci. **772:**140–151.
50. **Ladner, R. D., S. A. Carr, M. J. Huddleston, D. E. McNulty, and S. J. Caradonna.** 1996. Identification of a consensus cyclin-dependent kinase phosphorylation site unique to the nuclear form of human deoxyuridine triphosphate nucleotidohydrolase. J. Biol. Chem. **271:**7752–7757.
51. **Ladner, R. D., D. E. McNulty, S. A. Carr, G. D. Roberts, and S. J. Caradonna.** 1996. Characterization of distinct nuclear and mitochondrial forms of human deoxyuridine triphosphate nucleotidohydrolase. J. Biol. Chem. **271:** 7745–7751.
52. **Larsson, G., L. A. Svensson, and P. O. Nyman.** 1996. Crystal structure of the *Escherichia coli* dUTPase in complex with a substrate analogue (dUDP). Nature Struct. Biol. **3:**532–538.
53. **Lerner, D. A., P. C. Wagaman, T. R. Phillips, O. Prospero-Garcia, S. J. Henriksen, H. S. Fox, F. E. Bloom, and J. H. Elder.** 1995. Increased mutation frequency of feline immunodeficiency virus lacking functional deoxyuridine-triphosphatase. Proc. Natl. Acad. Sci. USA **92:**7480–7484.
54. **Lesk, A. M., and C. Chothia.** 1980. How different amino acid sequences determine similar protein structures: the structure and evolutionary dynamics of the globins. J. Mol. Biol. **136:**225–270.
55. **Matsushime, H., M. F. Roussel, R. A. Ashmun, and C. J. Sherr.** 1991. Colony-stimulating factor 1 regulates novel cyclins during the G1 phase of the cell cycle. Cell **65:**701–713.
56. **McAllister, R. M., M. B. Gardner, A. E. Greene, C. Bradt, W. W. Nichols, and B. H. Landing.** 1971. Cultivation *in vitro* of cells derived from a human osteosarcoma. Cancer **27:**397–402.
57. **McGeoch, D. J.** 1990. Protein sequence comparisons show that the 'pseudoproteases' encoded by poxviruses and certain retroviruses belong to the deoxyuridine triphosphatase family. Nucleic Acids Res. **18:**4105–4110.
58. **McIntosh, E. M., D. D. Ager, M. H. Gadsden, and R. H. Haynes.** 1992. Human dUTP pyrophosphatase: cDNA sequence and potential biological importance of the enzyme. Proc. Natl. Acad. Sci. USA **89:**8020–8024.
59. **McIntyre, M. C., M. G. Frattini, S. R. Grossman, and L. A. Laimins.** 1993. Human papillomavirus type 18 E7 protein requires intact Cys-X-X-Cys motifs for zinc binding, dimerization, and transformation but not for Rb binding. J. Virol. **67:**3142–3150.
60. **McKenzie, H. A., and F. H. White, Jr.** 1991. Lysozyme and α-lactalbumin: structure, function, and inter-relationships. Adv. Protein Chem. **41:**173–315.
61. **Meyer, A. N., Y. F. Xu, M. K. Webster, A. E. Smith, and D. J. Donoghue.** 1994. Cellular transformation by a transmembrane peptide: structural requirements for the bovine papillomavirus E5 oncoprotein. Proc. Natl. Acad. Sci. USA **91:**4634–4638.
62. **Moller, W., and R. Amons.** 1985. Phosphate-binding sequences in nucleotide-binding proteins. FEBS Lett. **186:**1–7.
63. **Myers, R., R. Williams, and R. Tjian.** 1981. Oligomeric structure of an SV40 T antigen in free form and as bound to SV40 DNA. J. Mol. Biol. **148:**347–353.
64. **Nilson, L. A., R. L. Gottlieb, G. W. Polack, and D. DiMaio.** 1995. Mutational analysis of the interaction between the bovine papillomavirus E5 transforming protein and the endogenous β receptor for platelet-derived growth factor in mouse C127 cells. J. Virol. **69:**5869–5874.
65. **Pang, J. H., and K. Y. Chen.** 1994. Global change of gene expression at late G1/S boundary may occur in human IMR-90 diploid fibroblasts during senescence. J. Cell. Physiol. **160:**531–538.
66. **Pardo, E. G., and C. Gutierrez.** 1990. Cell cycle- and differentiation stage-dependent variation of dUTPase activity in higher plant cells. Exp. Cell Res. **186:**90–98.
67. **Petti, L., and D. DiMaio.** 1992. Stable association between the bovine papillomavirus E5 transforming protein and activated platelet-derived growth factor receptor in transformed mouse cells. Proc. Natl. Acad. Sci. USA **89:** 6736–6740.
68. **Pyles, R. B., N. M. Sawtell, and R. L. Thompson.** 1992. Herpes simplex virus type 1 dUTPase mutants are attenuated for neurovirulence, neuroinvasiveness, and reactivation from latency. J. Virol. **66:**6706–6713.
69. **Qasba, P. K., and K. S. Surinder.** 1984. Similarity of the nucleotide sequences of rat α-lactalbumin and chicken lysozyme genes. Nature **308:**377–380.
70. **Reardon, D., and G. K. Farber.** 1995. The structure and evolution of alpha/beta barrel proteins. FASEB J. **9:**497–503.
71. **Richards, R. G., L. C. Sowers, J. Laszlo, and W. D. Sedwick.** 1986. The occurrence and consequences of deoxyuridine in DNA. Adv. Enzyme Regul. **22:**157–185.
72. **Roizman, B., and A. E. Sears.** 1990. Herpes simplex viruses and their replication, p. 1795–1841. *In* B. N. Fields, D. M. Knipe, et al. (ed.), Virology, 2nd ed., vol. 2. Raven Press, Ltd., New York, N.Y.
73. **Rost, B., C. Sander, and R. Schneider.** 1994. PHD—an automatic mail server for protein secondary structure prediction. Comput. Appl. Biosci. **10:**53–60.
74. **Sambrook, J., M. Sleigh, J. A. Engler, and T. R. Broker.** 1980. The evolution of the adenoviral genome. Ann. N. Y. Acad. Sci. **354:**426–452.
75. **Senkevich, T. G., J. J. Bugert, J. R. Sisler, E. V. Koonin, G. Darai, and B. Moss.** 1996. Genome sequence of a human tumorigenic poxvirus: prediction of specific host response-evasion genes. Science **273:**813–816.
76. **Smith, D. B., and L. M. Corcoran.** 1994. Expression and purification of glutathione-*S*-transferase fusion proteins, p. 16.7.1–16.7.7. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology. Greene Publishing Associates and Wiley-Interscience, New York, N.Y.
77. **Smith, R. F., and T. F. Smith.** 1990. Automatic generation of primary sequence patterns from sets of related protein sequences. Proc. Natl. Acad. Sci. USA **87:**118–122.
78. **Smith, R. F., and T. F. Smith.** 1992. Pattern-Induced Multi-sequence Alignment (PIMA) algorithm employing secondary structure-dependent gap penalties for comparative protein modelling. Protein Eng. **5:**35–41.
79. **Solovyev, V. V., and A. A. Salamov.** 1994. Predicting alpha-helix and beta-strand segments of globular proteins. Comput. Appl. Biosci. **10:**661–669.
80. **Steagall, W. K., M. D. Robek, S. T. Perry, R. J. Fuller, and S. L. Payne.** 1995. Incorporation of uracil into viral DNA correlates with reduced replication of EIAV in macrophages. Virology **210:**302–313.
81. **Strahler, J. R., X. X. Zhu, N. Hora, Y. K. Wang, P. C. Andrews, N. A. Roseman, J. V. Neel, L. Turka, and S. M. Hanash.** 1993. Maturation stage and proliferation-dependent expression of dUTPase in human T cells. Proc. Natl. Acad. Sci. USA **90:**4991–4995.
82. **Stuart, D. I., K. R. Acharya, N. P. C. Walker, S. G. Smith, M. Lewis, and D. C. Phillips.** 1986. Alpha-lactalbumin possesses a novel calcium binding loop. Nature **324:**84–87.
83. **Swofford, D. L.** 1993. PAUP: phylogenetic analysis using parsimony, version 3.1. Computer program distributed by the Illinois Natural History Survey, Champaign, Ill.
84. **Threadgill, D. S., W. K. Steagall, M. T. Flaherty, F. J. Fuller, S. T. Perry, K. E. Rushlow, G. S. Le, and S. L. Payne.** 1993. Characterization of equine infectious anemia virus dUTPase: growth properties of a dUTPase-deficient mutant. J. Virol. **67:**2592–2600.
85. **Trentin, J. J., Y. Yabe, and G. Taylor.** 1962. Quest for human cancer viruses. Science **137:**835–841.
86. **Turelli, P., G. Petursson, F. Guiguen, J. F. Mornex, R. Vigne, and G. Querat.** 1996. Replication properties of dUTPase-deficient mutants of caprine and ovine lentiviruses. J. Virol. **70:**1213–1217.
87. **Wagaman, P. C., L. C. Hasselkus, M. Henson, D. L. Lerner, T. R. Phillips, and J. H. Elder.** 1993. Molecular cloning and characterization of deoxyuridine triphosphatase from feline immunodeficiency virus (FIV). Virology **196:**451–457.
88. **Weiss, R. S., M. J. McArthur, and R. T. Javier.** 1996. Human adenovirus type 9 E4 open reading frame 1 encodes a cytoplasmic transforming protein capable of increasing the oncogenicity of CREF cells. J. Virol. **70:**862–872.
89. **Williams, M. V., and Y. Cheng.** 1979. Human deoxyuridine triphosphate nucleotidohydrolase. J. Biol. Chem. **254:**2897–2901.
90. **Xiong, Y., T. Connolly, B. Futcher, and D. Beach.** 1991. Human D-type cyclin. Cell **65:**691–699.
91. **Yeh, P., J. F. Dedieu, C. Orsini, E. Vigne, P. Denefle, and M. Perricaudet.** 1996. Efficient dual transcomplementation of adenovirus E1 and E4 regions from a 293-derived cell line expressing a minimal E4 functional unit. J. Virol. **70:**559–565.
92. **Yi, T. M., and E. S. Lander.** 1993. Protein secondary structure prediction using nearest-neighbor methods. J. Mol. Biol. **232:**1117–1129.