# Plant orthologs of p300/CBP: conservation of a core domain in metazoan p300/CBP acetyltransferase-related proteins

**Lorenza Bordoli, Marco Netsch, Urs Lüthi, Werner Lutz and Richard Eckner\***

Institute for Molecular Biology, University of Zurich, Winterthurerstraße 190, CH-8057 Zurich, Switzerland

## ABSTRACT

**p300 and CBP participate as transcriptional coregulators in the execution of a wide spectrum of cellular gene expression programs controlling cell differentiation, growth and homeostasis. Both proteins act together with sequence-specific transcription factors to modify chromatin structure of target genes via their intrinsic acetyltransferase activity directed towards core histones and some transcription factors. So far, p300-related proteins have been described in animals ranging from *Drosophila* and *Caenorhabditis elegans* to humans. In this report, we describe p300/CBP-like polypeptides in the plant *Arabidopsis thaliana*. Interestingly, homology between animal and plant p300/CBP is largely restricted to a C-terminal segment, about 600 amino acids in length, which encompasses acetyltransferase and E1A-binding domains. We have examined whether this conservation in sequence is paralleled by a conservation in function. The same amino acid residues critical for acetyltransferase activity in human p300 are also critical for the function of one of the plant orthologs. Remarkably, plant proteins bind to the adenovirus E1A protein in a manner recapitulating the binding specificity of mammalian p300/CBP. The striking conservation of an extended segment of p300/CBP suggests that it may constitute a functional entity fulfilling functions that may be essential for all metazoan organisms.**

## INTRODUCTION

p300 and CBP are large nuclear transcriptional regulatory proteins that are highly related in sequence and function. p300 was first to be identified as one of several cellular polypeptides binding to the adenoviral oncoprotein E1A (1,2). Since specific E1A mutants unable to bind p300 have lost the ability to stimulate growth and perturb cell differentiation, p300 has been linked to these processes (1,3). CBP, in turn, has been identified as a coactivator for the cAMP-responsive transcription factor CREB (4).

p300 and CBP are endowed with acetyltransferase activity (5,6) and interact, in distinct complexes, with numerous transcription factors through several evolutionarily conserved domains (7–10). In this way, acetyltransferase activity is thought to be recruited to specific promoters and enhancers where acetylation of neighboring nucleosomal histones facilitates access of other components required for transcription. In general, a positive correlation exists between the histone acetylation state of a gene and its transcriptional activity (11–14). However, in some cases, acetylation can also be associated with repressed chromatin. Acetylation of lysine 12 of histone H4 is specifically found in transcriptionally silenced chromatin in yeast (15).

Although histone acetylation is a common feature of eukaryotes (16–18), p300/CBP-related proteins so far appear to be restricted to multicellular organisms (higher eukaryotes) and cannot be found in the yeast *Saccharomyces cerevisiae* (19). This observation suggests that p300/CBP function in pathways specific to metazoan organisms such as cell-to-cell communication in organ development and morphogenesis. This view is supported by the phenotype of mice and flies carrying mutations in their p300 and CBP genes. Mice lacking either p300 or CBP die *in utero* between days 8.5 and 11.5 of gestation and invariably exhibit an open neural tube and frequently heart defects (20,21). In flies, only one p300/CBP-like gene has been described and mutations in it lead to embryonic lethality. *Drosophila* embryos harboring a hypomorphic dCBP allele display a twisted form and show altered expression of genes required for multiple developmental processes (22,23).

In this report, we describe plant (*Arabidopsis thaliana*) orthologs of p300 and CBP that were named p300/CBP acetyltransferase-related protein 1–4 (PCAT1–4). These plant proteins show a striking conservation in sequence of a region encompassing cysteine/histidine-rich domain 2 and 3 (C/H 2 and 3) in human p300/CBP, but they lack bromo- or CREB-binding (KIX) domains. In histone acetyltransferase (HAT) assays, PCAT2 but not PCAT1 is active. Surprisingly, despite the evolutionary distance to their human counterparts, PCAT1 and 2 bind to the adenovirus E1A protein with a specificity that matches that of human p300/CBP. Based on this remarkable

conservation, we propose that acetyltransferase (AT) and E1A binding domains form a unit that likely fulfills essential functions in all metazoan organisms.

## MATERIALS AND METHODS

### GenBank accession numbers

h*p300* complete coding sequence, accession no. U01877; h*CBP* complete coding sequence, accession no. U85962.3; *Arabidopsis* F1N21.18 (AthPCAT1), protein accession no. AAB95246.1; Ath*PCAT2* complete coding sequence, accession no. AF323954; F14J16.27 (AthPCAT3), protein accession no. AAF79331; *Arabidopsis* 'putative CREB-binding protein' (AthPCAT4), protein accession no. AAF35947.

### Identification of PCAT1 and 2

cDNAs for AthPCAT1 to 4 were identified by searching protein databases with the BEAUTY (24) (BLAST enhanced alignment utility) database search tool with the protein sequence of the AT domain of human p300. Multiple sequence analysis was performed with the clustalx software (25) and shading was done with the Boxshade program (http://ulrec3.unil.ch:80/software/BOX_form.html).

### Primers

See Table 1 for details of primers used.

### Cloning of PCAT1 and 2

RT–PCR (Titan One Tube RT–PCR System, Boehringer Mannheim) of segments of AthPCAT1 and AthPCAT2 messages was performed with total RNA isolated from *Arabidopsis* seedlings. For AthPCAT2, the full coding sequence was amplified and cloned. Sequencing showed that several exons were incorrectly predicted during the genome sequencing project. The corrected PCAT2 sequence has the GenBank accession no AF323954. For AthPCAT1, all of the coding sequence except the repeat motifs could be amplified and cloned. Three exons turned out to be incorrectly predicted in GenBank accession no AC002130.1. The three revised exons span nucleotides 109 595–109 670, 109 803–109 884 and 110 040–110 172 referring to the numbering used in AC002130.1.

The AT domains of AthPCAT1 containing amino acids 1055–1624 and of AthPCAT2 containing amino acids 985–1530 were each amplified as two fragments by RT–PCR, the two segments were combined and cloned into pGEX-KG (26). All cloned RT–PCR products were sequenced. The 5′ primers were 5′-PCAT1C (priming in exon 12) and 5′-PCAT1B (exon 8) for PCAT1, and 5′-PCAT2B (exon 10) and 5′-PCAT2A (exon 7) for PCAT2. The 3′ primers were 3′-PCAT1C (exon 16) and 3′-PCAT1B (exon 8) for PCAT1, and 3′-PCAT2B (exon 15) and 3′-PCAT2A (exon 7) for PCAT2.

For E1A binding studies, the C/H3 domains of AthPCAT1 containing amino acids 1588–1745 and of AthPCAT2

**Table 1.** List of primers used for RT–PCR amplification of PCATI–4 transcripts and for construction of GAL4- and GST fusion proteins

| Primer name | Sequence | Amplified fragment (amino acids) |
| --- | --- | --- |
| 5′-PCAT1A | 5′-CCGGGGTACCGAGGAAGACATGGATCTAAC-3′ | 1–367 |
| 3′-PCAT1A | 5′-GTTGAATACCTAGAGGGCCCATCTC-3′ | |
| 5′-PCAT1B | 5′-TCCCCCCGGGGCCAAGTACACAGAGCGCAGTC-3′ | 1013–1318 |
| 3′-PCAT1B | 5′-CCAATCAATACTTCATGATAAACG-3′ | |
| 5′-PCAT1C | 5′-CGTTTATCATGAAGTATTGATTGG-3′ | 1311–1736 |
| 3′-PCAT1C | 5′-GGAGGAGCTCGAGATCTTGGCAATGGTACAC-3′ | |
| 5′-PCAT1D | 5′-TCCGGCTAGCTTCGATGATGATACTTCACCATC-3′ | 1594–1745 |
| 3′-PCAT1D | 5′-GGAGGAGCTCTTACCTGCACTGCGGTACGGGGCAATTGAGATCTTGG-3′ | |
| 5′-PCAT2A | 5′-CGACTCTAGACTTTTGTATTCCATGTTATAATGAATCC-3′ | 985–1213 |
| 3′-PCAT2A | 5′-CTCCGTTATAAGATCTAACCTCGG-3′ | |
| 5′-PCAT2B | 5′-GACCCGAGGTTAGATCTTATAACGGAG-3′ | 1206–1530 |
| 3′-PCAT2B | 5′-CCATGAGCTCGCCTTGACCAGTTTCAATGTCAA-3′ | |
| 5′-PCAT2C | 5′-GTCATCTAGAGTCCATGATGGTCCTTTATCATCT-3′ | 1499–1655 |
| 3′-PCAT2C | 5′-GTCCTTCGAATTACCTGCAACGAGGTACATGGC-3′ | |
| 5′-PCAT3A | 5′-CCTCCACTTAAGGGTGACGA-3′ | 1081–1257 |
| 3′-PCAT3A | 5′-GCTTGTCACAGATCTGAAAATTC-3′ | |
| 5′-PCAT4A | 5′-CAGAGTTGTAGCATCTGTCG | 1111–1456 |
| 3′-PCAT4A | 5′-CTCGCTCTCGAGGATGTCATCATTG-3′ | |
| 5′-PCAT2–GAL4 | 5′-GACGTCTAGATCGTCATTACTTTTGTATTCCATGTTATAATG-3′ | 982–1530 |
| 3′-PCAT2–GAL4 | 5′-CTAGTCTAGACTACAGGTCCTCCTCGGAGATCAGCTTCTGCTCCAGGTCCTCCTCGGAGATCAG-CTTCTGCTCGCCTTGACCAGTTTCAATG-3′ | |

containing amino acids 1499–1655 were amplified by RT–PCR and cloned into pGEX-KG. The 5′ primers were 5′-PCAT1D (exon 16) for PCAT1 and 5′-PCAT2C (exon 15) for PCAT2. The reverse transcription primers were 3′-PCAT1D (exon 16) for PCAT1 and 3′-PCAT2C (exon 16) for PCAT2.

The GAL4–PCAT2 HAT fusion constructs were made by cloning the PCR-amplified wild-type or WY-mutant HAT domain of PCAT2 (982–1530) into the *Xba*I restriction site of pcDNA3 vector (Stratagene) containing the GAL4–DNA binding domain (DBD) (1–147) (27). PCR primers used were 5′-PCAT2–GAL4 and 3′-PCAT2–GAL4.

### Site-directed mutagenesis

The QuickChange Site-Directed Mutagenesis Kit from Stratagene was used to introduce the mutations FPY(amino acids 1353–1355)AAA, WY(amino acids 1466–1467)AS and ML(amino acids 1470–1471)AA into the AT domain of hp300. Mutations FAY(amino acids 1153–1155)AAA, WY(amino acids 1273–1274)AA and ML(amino acids 1277–1278)AA were introduced into the AT domain of AthPCAT2. The mutation WY(amino acids 1503–1504)AS was also introduced into the AT domain of mCBP.

### Northern blot analysis

Total RNA extracted from different tissues of *A.thaliana* line (L.) Heynh. var. Columbia was kindly provided by Dr Kay Schneitz and Patrick Sieber (Institute for Plant Biology, University of Zurich). RNA from flowers was isolated from flower stages 1–13 [stages according to Smyth *et al.* (28)]. Ten micrograms of RNA per sample was fractionated on a denaturing 0.8% agarose–formaldehyde gel (29) and transferred in 50 mM NaOH by capillary blotting onto Hybond N+ nylon membrane (Amersham). Hybridization was performed with a 1.0 kb Ath*PCAT2* cDNA fragment labeled with $^{32}$P by random priming (Prime-It II Random Primer Labeling Kit, Stratagene) in the presence of 50% formamide at 42°C for 24 h. The membrane was washed at high stringency with 0.1× SSC, 0.1% SDS at 65°C and hybridization signals were analyzed with a PhosphoImager (Molecular Dynamics).

### RT–PCR analysis of PCAT1–4 mRNA expression

RT–PCR (Titan One Tube RT–PCR System, Boehringer Mannheim) of AthPCAT1–4 transcript levels was performed with total RNA isolated from the indicated tissues of *A.thaliana*.

The 5′ primers were 5′-PCAT1A for PCAT1, 5′-PCAT2A for PCAT2, 5′-PCAT3A for PCAT3 and 5′-PCAT4A for PCAT4. The reverse transcription primers were 3′-PCAT1A for PCAT1, 3′-PCAT2A for PCAT2, 3′-PCAT3A for PCAT3 and 3′-PCAT4A for PCAT4. As a negative control, 10 μg of total yeast RNA was amplified with each PCAT primer pair and as a positive control a region of the ubiquitously expressed GapC mRNA (spanning amino acids 129–309) was amplified.

### Histone acetyltransferase assay

For bacterial expression, cDNA fragments corresponding to the catalytic domain of hp300 (amino acids 1197–1703), AthPCAT1 (amino acids 1055–1624) and AthPCAT2 (amino acids 985–1530) were cloned into the pGEX–KG vector (26), expressed in *Escherichia coli* and the glutathione *S*-transferase (GST)-fusion proteins were purified with glutathione–Sepharose (Pharmacia). HAT assays were performed with the recombinant proteins using a modification of the assay previously described (30). Enzymes samples were incubated at room temperature for 60 min in 30 μl of assay buffer [50 mM Tris–HCl pH 8.0, 10% (v/v) glycerol, 1 mM DTT, 1 mM PMSF, 10 mM sodiumbutyrate] containing 0.25 nCi of [$^{14}$C]acetyl-CoA (55mCi/mmol, Hartmann) and 5 μg calf thymus histones (Sigma, type IIA). The reaction products were resolved on SDS–polyacrylamide gel, which was treated with Na-salycilate for fluorography.

### Cell culture, transfections and reporter gene assay

U2OS cells were maintained in DMEM supplemented with 10% fetal calf serum (ICN). Transient transfection assays were carried out with the calcium phosphate precipitation method (31) using 3 μg of CMV-promoter driven GAL4–CBP HAT (27) or GAL4–PCAT2 HAT (wild-type or the WY mutant version) together with 10 μg of 5× GAL4–AdML-luciferase reporter and 1 μg of CMV–β-galactosidase plasmid as internal standard. The following day, cells were washed and were harvested 36 h after transfection. Cells were resuspended in 100 μl extraction buffer (100 mM potassium phosphate pH 7.8, 1 mM DTT) and lysed by freezing and thawing three times. Supernatants were clarified by centrifugation (5 min, 12 000 *g*), 10 μl of extract were mixed with 100 μl luciferase assay solution (Promega) and analyzed in a luminometer (EG&G Berthold Lumat LB 9507). β-galactosidase assay was performed according to standard methods using 50 μl of the extract and luciferase units were normalized according to β-galactosidase values.

### E1A binding assay

U2OS cells were infected with adenovirus type 5 (Ad5 wt) expressing wild-type 12S E1A protein, or the mutant viruses dl 312 (32), expressing no E1A protein due to a large deletion in the early region, dl 1101 expressing E1A harboring a deletion in the N-terminal E1A binding domain (amino acids 5–24), or dl 1107 expressing an E1A protein lacking most of conserved region 2 (amino acids 111–123) (33–35) at a multiplicity of infection of 10. U2OS cells were harvested 12–15 h after infection, lysed in EBC (50 mM Tris–HCl pH 8.0, 250 mM NaCl, 0.5% NP-40) and cell lysates were incubated with purified GST–C/H3 domain fusion proteins containing the C/H3 domain of hp300, AthPCAT1 or AthPCAT2 at 4°C for 1 h. Subsequently, binding reactions were washed five times with EBC, resuspended in SDS sample buffer and the bound E1A was detected in western blots with the E1A monoclonal antibody M73 (36).

## RESULTS

Protein sequence database searches with the peptide sequence of the C/H 2 of human p300 led to the identification of four *A.thaliana* proteins with striking homologies to human p300 and CBP. Closer inspection showed that sequence similarities are mainly clustered in a 600 amino acid segment encompassing C/H 2 and 3 regions of mammalian p300/CBP. This region harbors AT and E1A-binding domains. Based on this remarkable conservation, the plant proteins were named p300/CBP acetyltransferase-related protein 1–4 (PCAT1–4). Figure 1A schematically depicts regions of highest homology as black boxes and Figure 1B shows an alignment of the C-terminal
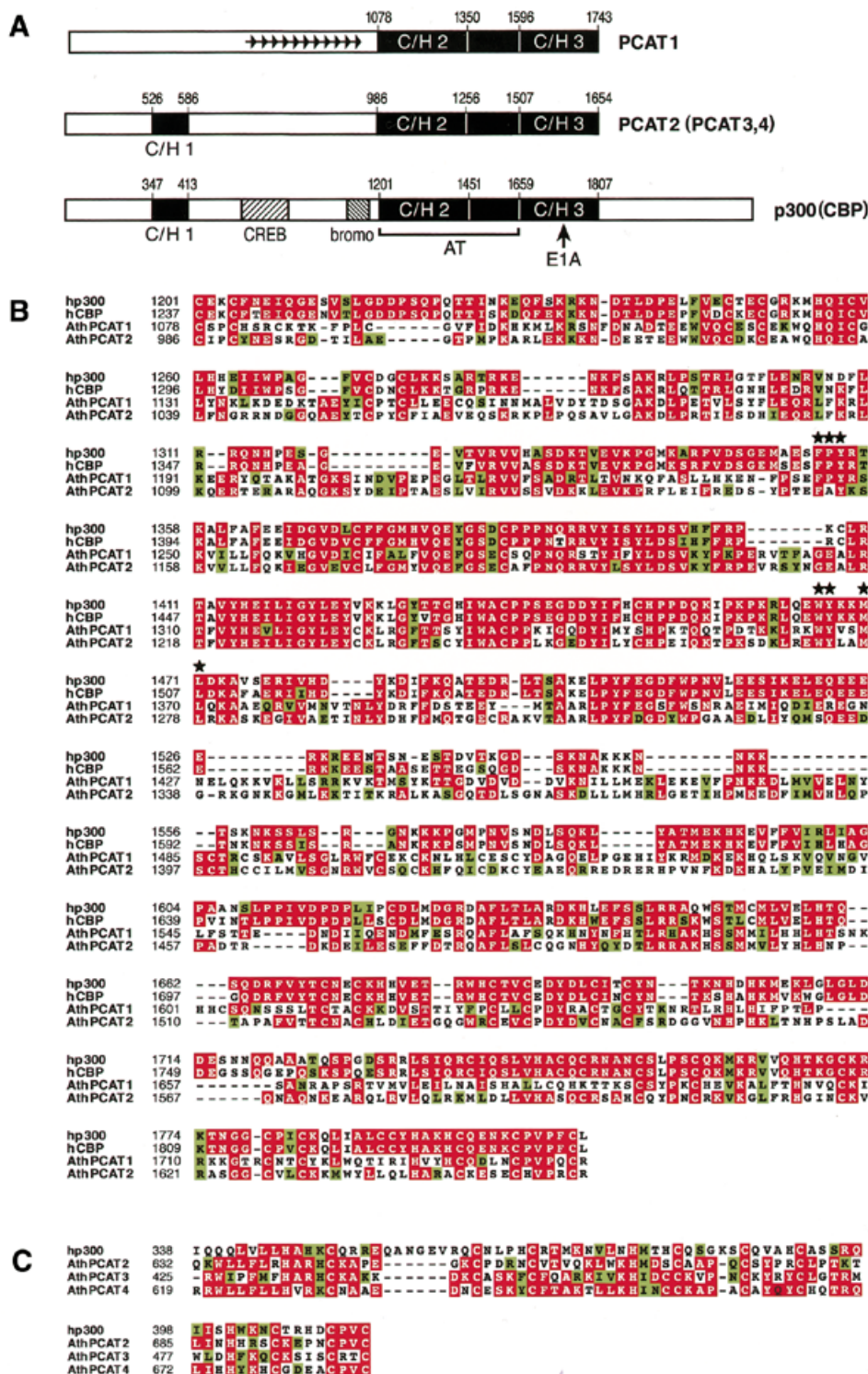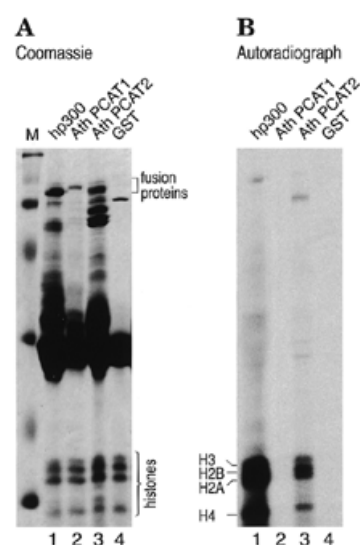
**Figure 1.** Identification of p300/CBP-like molecules in the plant *A.thaliana*. (**A**) Schematic representations of human p300 (CBP) and PCAT1 and 2 are shown. PCAT3 and 4 are not depicted because they structurally resemble PCAT2. Regions of highest homology are highlighted as black boxes and the amino acids located at the borders of the C/H regions are numbered. Note the absence of the CREB binding region (marked CREB) and a bromodomain (marked bromo) in the plant PCAT proteins. (**B**) Alignment of the segments encompassing C/H2 and C/H3 regions of human p300, CBP and the two plant proteins PCAT1 and 2. The shading of the multiple sequence alignment was done using the standard settings of the Boxshade 3.21 program. Amino acids are shaded when proteins exhibit similar (green) or identical (red) residues at the respective position according to the rules specified in the documentation to the algorithm of the software: http://www.ch.embnet.org/software/BOX_doc.html. Asterisks above the aligned sequences indicate AT domain residues that were mutated in p300 and PCAT2 in Figure 4. (**C**) A sequence alignment of the C/H1 regions of hp300 and AthPCAT2, 3 and 4 is shown.
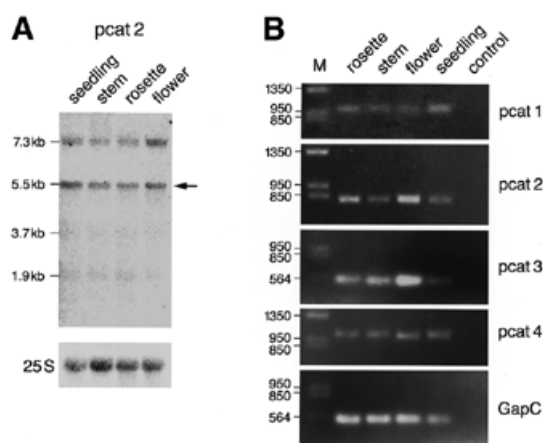
**Figure 2.** Ubiquitous expression of PCAT1–4 mRNAs. (**A**) Northern blot using total RNA isolated from *A.thaliana* seedling, stem, rosette or flower tissue. The blot was probed with a PCAT2 fragment and shows two hybridizing bands of ~5.5 and 7.3 kb. Both mRNAs are of sufficient length to harbor the complete PCAT2 protein coding region, which is 4.7 kb in length. The 5.5 kb band is marked by an arrow because it is likely to represent the mature form of the PCAT2 transcript. The panel at the bottom shows the hybridization signal of the 25S ribosomal RNA serving as a loading control. (**B**) PCAT1, 2, 3 and 4 transcripts are expressed in early stages of development (seedlings) and in the three major tissues (rosette, stem and flowers) of the adult plant. Fragments of each of the four PCAT transcripts (see Table 1 for location of amplified fragments) were amplified by RT–PCR from total RNA isolated from the indicated tissue. Amplification was stopped while it was still in the linear range. Signals should not be taken to reflect the precise relative or absolute levels of mRNA abundance.



**Figure 3.** PCAT2 exhibits HAT activity. GST-fusion proteins containing the AT domain of AthPCAT1 (amino acids 1055–1624) and AthPCAT2 (amino acids 985–1530) were incubated with calf thymus histones and [$^{14}$C]acetyl-CoA and the proteins were analyzed by SDS−PAGE followed by Coomassie blue staining [shown in (**A**)] and viewed following autoradiography of the gel (**B**). For a comparison of relative activity, the AT domain of human p300 was expressed in parallel [lane 1 in (A) and (B)]. Of the two plant proteins, only the AT domain of PCAT2 exhibited enzymatic activity, suggesting that at least PCAT2 is indeed a bona fide HAT. Even after prolonged exposure, no histone acetylation could be detected in the PCAT1 lane.

homology segment of human p300 and the two plant proteins PCAT1 and 2 (33 and 35% identity, respectively). It will be demonstrated below that at least PCAT2 indeed possesses acetyltransferase activity. Outside this segment, no significant homologies are present, with the exception of a stretch of about 60 amino acids near the N-terminus of PCAT2, 3 and 4 exhibiting homology to the first cysteine/histidine-rich region of p300/CBP (Fig. 1C). Interestingly, none of the PCAT proteins contains a bromodomain and a CREB-binding (KIX) region, both of which are highly conserved among animal p300/CBP proteins. Instead, PCAT1 harbors 11 copies of a repetitive motif that does not display sequence similarity to any other motif.

Since PCAT1 and 2 protein sequences were derived from hypothetical transcripts assembled by computer from genomic DNA fragments sequenced in the course of the *A.thaliana* genome sequencing project, we wanted to ascertain that mRNAs with the structure of the predicted transcripts indeed existed. A northern blot of total RNA isolated from seedlings, stem, rosette and flowers probed with a PCAT2 fragment showed two hybridizing transcripts of ~5.5 and 7.3 kb (Fig. 2A). Both mRNAs are of sufficient length to harbor the complete PCAT2 protein-coding region (4.7 kb). The two transcripts may represent alternatively spliced forms of the PCAT2 mRNA or the larger transcript could be a precursor transcript to the smaller one. To assay for expression of the other PCAT transcripts, selected fragments of all four PCAT mRNAs were individually amplified by RT–PCR approach. PCR products were cloned and sequenced. In both PCAT1 and 2 genes, some

exons turned out to be incorrectly predicted based on the genomic sequence (see GenBank accession no AF323954 for PCAT2 and Materials and Methods for PCAT1). Similar to their mammalian orthologs, all four PCAT mRNAs are expressed in early stages of development (seedlings) and in the three major tissues (rosette, stem and flowers) of adult *Arabidopsis* plants (Fig. 2B). As a negative control, yeast RNA was amplified with each PCAT primer pair (lane labeled control) and the ubiquitously expressed cytosolic glyceraldehyde-3-phosphate-dehydrogenase (GapC) mRNA (37) was amplified as a positive control (Fig. 2B, bottom panel).

We next examined whether the region of PCAT1 and 2 homologous to the AT domain of hp300/CBP would also be enzymatically active. For this purpose, the AT domain of the two plant proteins was cloned and expressed in *E.coli* as a GST fusion protein. For a comparison of relative activity, the AT domain of human p300 was expressed in parallel. Standard HAT assays were performed using core histones as substrate. Figure 3A shows a protein gel stained with Coomassie blue revealing expression levels of GST–AT fusion proteins and Figure 3B shows an autoradiograph of the same protein gel. The AT domain of PCAT1 turned out to be inactive in this assay (Fig. 3B, lane 2), while the same domain of PCAT2 and human p300 exhibited clearly detectable HAT activity (Fig. 3B, lanes 1 and 3, respectively). The AT domain of PCAT1 was also inactive when produced in a plant environment by *in vitro* translation in a wheat germ extract. In the same experiment, the AT domain of human p300 was active (data not shown). We conclude that at least PCAT2 is a bona fide HAT.
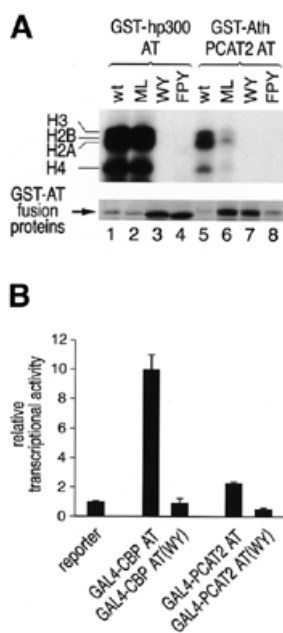
**Figure 4.** Conserved residues which are essential for the AT activity of human p300 and CBP are also essential for the enzymatic activity of *Arabidopsis* PCAT2 and for stimulation of transcription by the respective AT domains. (**A**) Specific point mutations were introduced into the GST–AT domain fusion proteins of both species (see Fig. 1B and Materials and Methods for exact location of mutations) and HAT assays were performed. The top panel shows an autoradiograph of the SDS–PAGE gel demonstrating that mutation of FPY or WY residues abolished AT activity of both human p300 and plant PCAT2. By contrast, mutation of the residues ML did not abolish AT activity of hp300 and reduced activity of the plant protein by a factor of about 3. The bottom panel shows a Coomassie blue stained section of the same protein gel revealing expression levels of the hp300 and PCAT2 GST–AT fusion proteins. (**B**) U2OS cells were transiently transfected with a GAL4-dependent reporter gene and GAL4–AT expression plasmids. GAL4–CBP AT strongly activated transcription of the reporter plasmid in a manner dependent on AT activity since GAL4–CBP (WY) was inactive. In comparison, GAL4–PCAT2 AT also activated transcription in an AT-dependent manner, but less efficently than the CBP AT domain.

To assess the relevance of the sequence similarity between the AT domains of human p300 and PCAT2 in closer detail, amino acid residues conserved between plant and human proteins were mutated to alanine residues at three different locations within the AT domains. The residues mutated (FPY in hp300, FAY in AthPCAT2; WY and ML in each hp300 and AthPCAT2, respectively) are indicated in Figure 1B by asterisks at the top of the alignment. Wild-type and mutant AT domains were again expressed in *E.coli* as GST-fusion proteins and a HAT assay was performed. Figure 4A demonstrates that mutation of FPY or WY residues abolished AT activity of both human p300 (lanes 3 and 4) and plant PCAT2 (lanes 7 and 8). By contrast, mutation of the residues ML did not affect enzymatic activity of the human p300 AT domain and reduced activity of the plant AT domain by a factor of three. Expression levels of the various GST–AT fusion proteins are shown at the bottom of Figure 4A. The result of our mutagenesis suggests that some of the conserved residues are essential for enzymatic activity of both human and plant proteins and emphasizes the significance of the observed sequence similarity.

The AT domain of CBP has been reported to activate transcription in mammalian cells when fused to the GAL4–DBD (27). In these experiments, activation of transcription was correlated with the enzymatic activity of the AT domain. To examine whether the PCAT2 AT domain is also able to stimulate transcription in a manner dependent on AT activity in mammalian cells, we cloned the wild-type and WY-mutated PCAT2 AT domain in-frame to the GAL4–DBD. Human U2OS cells were transiently transfected with a GAL4-dependent reporter plasmid and GAL4–AT expression plasmids. GAL4–CBP AT strongly activated transcription in a manner dependent on acetyltransferase activity since GAL4–CBP AT (WY) was inactive (Fig. 4B). In comparison, GAL4–PCAT2 AT also activated transcription in an acetyltransferase-dependent manner, but less efficiently than the CBP AT domain (Fig. 4B). Considering the results of Figure 4A and B together, the enzymatic activity of the AT domains measured *in vitro* correlates quite well with the transcriptional activity of these domains measured *in vivo*.

We next focused on the C/H3 region. Close inspection of the alignment of the C/H3 region revealed that this region is somewhat shorter than previously thought, based on sequence alignments of animal p300/CBP proteins. Some of the cysteine and histidine residues in the N-terminal part of the human p300 C/H3 domain are not conserved in the plant PCAT proteins, suggesting that the core of the C/H3 domain actually begins at hp300 residue 1659.

The C/H3 region has previously been characterized as a high-affinity binding site for the adenoviral E1A oncoprotein (2,38). To test whether the plant C/H3 region has the ability to interact in a manner similar to human p300 with E1A, we produced the C/H3 domain of human p300 and *Arabidopsis* PCAT1 and 2 as GST fusion proteins. The purified fusion proteins were incubated with extracts prepared from U2OS cells infected with adenoviruses expressing different E1A versions (see Fig. 5A for the structure of the expressed E1A proteins). The wild-type virus expresses wild-type 12S E1A, dl312 does not express any E1A due to a deletion including the E1A gene, dl1101 expresses a mutant form of E1A lacking amino acids 4–25 (constituting part of the p300 binding site) and dl1107 expresses E1A carrying a mutation in CR2 of E1A (constituting part of the RB protein family binding site). Remarkably, both plant GST–PCAT–C/H3 fusion proteins recapitulated the binding specificity of human p300 in that they bound to wild-type and dl1107 mutant E1A protein but not to the dl1101 mutant E1A (Fig. 5B and C, compare lanes 1–4 with lanes 9–12). These results provide evidence for the existence of a remarkably conserved protein binding surface in the C/H3 domain of human and plant p300/CBP-related proteins.

## DISCUSSION

So far, p300 and CBP-related proteins have been described in animals but not in plants. This paper closes a gap by characterizing in part plant proteins sharing several of the sequence motifs and functional properties of animal p300/CBP. Remarkably, our analysis points to a highly conserved core segment, encompassing C/H2 and C/H3 domains, which is present in all known p300/CBP proteins. Since unicellular eukaryotes such as the yeast *S.cerevisiae* do not contain p300/CBP-related
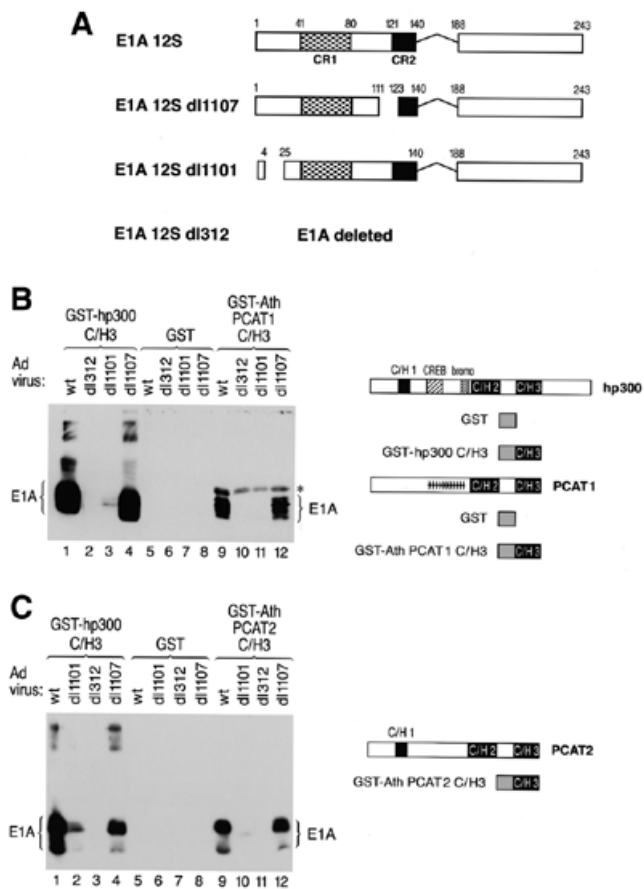
**Figure 5.** The C/H3 domain of PCAT1 and 2 specifically binds to E1A and discriminates between mutant E1A proteins in a manner recapitulating the binding pattern of human p300. (**A**) A schematic representation of the different adenovirus mutants used in the E1A binding assay is shown. The wild-type virus expresses wild-type 12S E1A, dl1107 expresses E1A carrying a mutation in conserved region 2 (CR2) of E1A (constituting part of the RB protein family binding site), dl1101 expresses a mutant form of E1A lacking amino acids 5–24 (constituting part of the p300 binding site) and dl312 does not express E1A at all due to a deletion in the early region removing the E1A gene. (**B** and **C**) Shown on the right are schematic representations of hp300 and PCAT1, 2 and the GST-C/H3 fusion proteins used for the E1A interaction assays. Purified GST proteins containing the C/H3 domain of hp300 [(B) and (C) lanes 1–4], of PCAT1 [(B) lanes 9–12], PCAT2 [(C) lanes 9–12] or GST alone [(B) and (C) lanes 5–8] were incubated with cell lysates prepared from U2OS cells infected with the indicated wild-type or mutant adenovirus. Bound E1A proteins were detected by western blot. The asterisk in (B) marks a non-specific background band visible in lanes 9–13. This band is due to binding of the secondary antibody to an *E.coli* protein copurifying with the GST−AthPCAT1−C/H3 fusion protein.

proteins (19), we propose that this core segment of p300/CBP fulfills essential functions in the context of processes unique to higher eukaryotes (metazoans) such as cell differentiation and organ development. Such a view is compatible with previous results demonstrating the importance of p300/CBP for animal development. Flies carrying a P element insertion in dCBP, the only known p300/CBP-related protein in *Drosophila melanogaster*, die at the earliest stages of embryogenesis (22,23) and mice lacking either p300 or CBP die around mid-gestation exhibiting a severe developmental retardation (20,21).

Interestingly, similar to mice and humans, more than one p300/CBP-like protein exists in the plant *A.thaliana*. While writing this manuscript, another two gene products sharing high sequence similarity with PCAT1 and 2 appeared in the database. These two proteins are referred to as PCAT3 and 4 and are more similar to PCAT2 than PCAT1 (see also Fig. 1). Considering the structural differences, especially between PCAT1 and 2, it is conceivable that the *in vivo* roles of the four AthPCAT proteins may not be entirely overlapping. PCAT2, 3 and 4 are most similar to human p300/CBP in that they possess all three regions rich in cysteine- and histidine residues, including a C/H1 domain (see Fig. 1C), while PCAT1 only harbors the segment including C/H2 and C/H3 and lacks a C/H1 domain. Also, on a functional level, PCAT2 appears to be more similar than PCAT1 to human p300/CBP in that it exhibits readily detectable HAT activity. This similarity is further underscored by the results of our substitution mutations of amino acid residues conserved between the AT domain of human p300/CBP and plant PCAT2. Both mutations abolishing enzymatic activity of the human protein also eliminate activity of the plant protein.

Unexpectedly, despite a comparable level of homology (~34% identity) between human p300 and plant PCAT1 and 2 AT domains, we were unable to detect any HAT activity associated with PCAT1 in several different *in vitro* assays. This negative result does not rule out that PCAT1 *in vivo* also exhibits acetyltransferase activity. For example, the recombinant PCAT1 fragment produced in *E.coli* or in wheat germ extracts may lack some modifications essential for its enzymatic activity. Alternatively, since our assays used histones as substrate, these proteins may not be the physiologically relevant *in vivo* substrate for PCAT1. Finally, it is possible that the AT domain harbors an important function other than acetyltransferase activity, which could explain sequence conservation.

In contrast to AT activity, both PCAT 1 and 2 were capable to interact in a specific manner via their C/H3 domain with the adenovirus E1A protein. This result suggests a remarkable conservation in metazoans of aspects of sequence, and maybe also structure, of the C/H3 domain. In mammals, many different factors are known to bind to the C/H3 domain (7–9). Some of the reported C/H3 interacting proteins exhibit enzymatic activity by themselves, such as the acetyltransferase PCAF (39), which also exists in plants, and the protein kinase RSK-1 (40). It is likely that cellular proteins conserved between animals and plants exist that bind to the C/H3 domain and exert essential functions.

The striking conservation of the C/H2 and C/H3 domains, including the intervening sequence, suggests that this segment may represent a functionally linked entity. The C/H3 region of this functional unit, as reported in a recent study (41), appears to be a critical element, significantly affecting the enzymatic activity of the neighboring core AT region under certain circumstances.

Remarkably, bromo and CREB binding (KIX) domains appear to be dispensable for the *Arabidopsis* p300/CBP orthologs. In addition, human p300 and CBP possess C-terminal to the C/H3 domain, an extended glutamine-rich C-terminus harboring the binding site for another group of acetyltransferases known as SRC-1 or ACTR. This extended C-terminus is completely absent in all four PCAT proteins.

The bromodomain has been shown to interact specifically with acetylated lysine and it has been suggested that such an interaction may facilitate binding of proteins containing bromodomains to transcriptionally active chromatin (42). It is intriguing that AthPCAT proteins can fulfill their functions without a bromodomain. We also note that other *Arabidopsis* proteins such as AthGCN5 do contain a bromodomain (43). The most well defined interaction partner for the KIX domain is the transcription factor CREB, which also appears to exist in at least some plants (44). Of potential interest is the occurrence of the short repeat motifs in PCAT1, which might substitute for one of the above mentioned motifs.

Since the first description of HAT activity in plant cells (45), multiple HATs have been characterized in plants, mostly in maize, and the majority belong to the category of cytoplasmatic HATs. However, members of the nuclear type of HATs have also been identified in plants, such as the above mentioned *Arabidopsis* GCN5 or the HATs involved in the embryo germination of *Zea mays* (46–48). This report shows that apart from the animal kingdom, the plant kingdom also expresses large proteins with several of the key features of animal p300/CBP proteins. It thus appears that p300/CBP-related proteins are common to all multicellular organisms, where they may be involved in functions specific to metazoans such as cell and organ differentiation processes, which require an intricate ability to coordinate cell-to-cell communications.

## ACKNOWLEDGEMENTS

## NOTE ADDED IN PROOF

The sequence of the *Arabidopsis thaliana* genome has recently been completed [The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, **408**, 796–815]. The authors of this paper have given the following names to the Arabidopsis p300/CBP orthologs: HAC1 (PACT2), HAC2 (PACT1), HAC4 (PACT3), HAC5 (PACT4). We recommend to apply the nomenclature of the Arabidopsis Genome Initiative (http://Ag.Arizona.Edu/chromatin/chromatin.html).

## REFERENCES

1. Harlow,E., Whyte,P., Franza,R.,Jr and Schley,C. (1986) Association of adenovirus early-region 1A proteins with cellular polypeptides. *Mol. Cell. Biol.*, **6**, 1579–1589.
2. Eckner,R., Ewen,M.E., Newsome,D., Gerdes,M., DeCaprio,J.A., Lawrence,J.B. and Livingston,D.M. (1994) Molecular cloning and functional analysis of the adenovirus E1A-associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor. *Genes Dev.*, **8**, 869–884.
3. Bayley,S.T. and Mymryk,J.S. (1994) Adenovirus E1A proteins and transformation. *Int. J. Oncol.*, **5**, 425–444.
4. Chrivia,J.C., Kwok,R.P.S., Lamb,N., Hagiwara,M., Montminy,M.R. and Goodman,R.H. (1993) Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature*, **365**, 855–859.
5. Bannister,A.J. and Kouzarides,T. (1996) The CBP co-activator is a histone acetyltransferase. *Nature*, **384**, 641–643.
6. Ogryzko,V.V., Schiltz,R.L., Russanova,V., Howard,B.H. and Nakatani,Y. (1996) The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell*, **87**, 953–959.
7. Shikama,N., Lyon,J. and LaThangue,N.B. (1997) The p300/CBP family: integrating signals with transcription factors and chromatin. *Trends Cell Biol.*, **7**, 230–236.
8. Giles,R.H., Peters,D.J.M. and Breuning,M.H. (1998) Conjunction dysfunction: CBP/p300 in human disease. *Trends Genet.*, **14**, 178–183.
9. Blobel,G.A. (2000) CREB-binding protein and p300: molecular integrators of hematopoietic transcription. *Blood*, **95**, 745–755.
10. Giordano,A. and Avataggiati,M.L. (1999) p300 and CBP: partners for life and death. *J. Cell Physiol.*, **181**, 218–230.
11. Kuo,M.-H., Zhou,J., Jambeck,P., Churchill,M.E.A. and Allis,C.D. (1998) Histone acetyltransferase activity of yeast Gcn5p is required for the activation of target genes in vivo. *Genes Dev.*, **12**, 627–639.
12. Wang,L., Liu,L. and Berger,S.L. (1998) Critical residues for histone acetylation by Gcn5, functioning in Ada and SAGA complexes, are also required for transcriptional function in vivo. *Genes Dev.*, **12**, 640–653.
13. Struhl,K. (1998) Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev.*, **12**, 599–606.
14. Strahl,B.D. and Allis,C.D. (2000) The language of covalent histone modifications. *Nature*, **403**, 41–45.
15. Braunstein,M., Sobel,R.E., Allis,C.D., Turner,B.M. and Broach,J.R. (1996) Efficient transcriptional silencing in Saccharomyces cerevisiae requires a heterochromatin histone acetylation pattern. *Mol. Cell. Biol.*, **16**, 4349–4356.
16. Grunstein,M. (1997) Histone acetylation in chromatin structure and transcription. *Nature*, **389**, 349–352.
17. Cheung,W.L., Briggs,S.D. and Allis,C.D. (2000) Acetylation and chromosomal functions. *Curr. Opin. Cell Biol.*, **12**, 326–333.
18. Sterner,D.E. and Berger,S.L. (2000) Acetylation of histones and transcription-related factors. *Microbiol. Mol. Biol. Rev.*, **64**, 435–459.
19. Eckner,R. (1996) p300 and CBP as transcriptional regulators and targets of oncogenic events. *Biol. Chem.*, **377**, 685–688.
20. Yao,T.-P., Oh,S.P., Fuchs,M., Zhou,N.-D., Ch'ng,L.-E., Newsome,D., Bronson,R.T., Livingston,D.M. and Eckner,R. (1998) Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. *Cell*, **93**, 361–372.
21. Kung,A.L., Rebel,V.I., Bronson,R.T., Ch'ng,L.E., Sieff,C.A., Livingston,D.M. and Yao,T.P. (2000) Gene dose-dependent control of hematopoiesis and hematologic tumor suppression by CBP. *Genes Dev.*, **14**, 272–277.
22. Akimaru,H., Hou,D.X. and Ishii,S. (1997) Drosophila CBP is required for dorsal-dependent twist gene expression. *Nature Genet.*, **17**, 211–214.
23. Akimaru,H., Chen,Y., Dai,P., Hou,D.X., Nonaka,M., Smolik,S.M., Armstrong,S., Goodman,R.H. and Ishii,S. (1997) Drosophila CBP is a co-activator of cubitus interruptus in hedgehog signalling. *Nature*, **386**, 735–738.
24. Worley,K.C., Wiese,B.A. and Smith,R.F. (1995) BEAUTY: an enhanced BLAST-based search tool that integrates multiple biological information resources into sequence similarity search results. *Genome Res.*, **5**, 173–184.
25. Thompson,J.D., Gibson,T.J., Plewniak,F., Jeanmougin,F. and Higgins,D.G. (1997) The CLUSTAL:X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.*, **25**, 4876–4882.
26. Guan,K.L. and Dixon,J.E. (1991) Eukaryotic proteins expressed in Escherichia coli: an improved thrombin cleavage and purification procedure of fusion proteins with glutathione S-transferase. *Anal. Biochem.*, **192**, 262–267.
27. Martinez-Balbas,M.A., Bannister,A.J., Martin,K., Seuffert,P.-H., Meisterernst,M. and Kouzarides,T. (1998) The acetyltransferase activity of CBP stimulates transcription. *EMBO J.*, **17**, 2886–2893.
28. Smyth,D.R., Bowman,J.L. and Meyerowitz,E.M. (1990) Early flower development in Arabidopsis. *Plant Cell*, **2**, 755–767.
29. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

30. Brownell,J.E. and Allis,C.D. (1995) An activity gel assay detects a single, catalytically active histone acetyltransferase subunit in Tetrahymena macronuclei. *Proc. Natl Acad. Sci. USA*, **92**, 6364–6368.
31. Chen,C. and Okayama,H. (1987) High-efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.*, **7**, 2745–2752.
32. Jones,N. and Shenk,T. (1979) An adenovirus type 5 early gene function regulates expression of other early viral genes. *Proc. Natl Acad. Sci. USA*, **76**, 3665–3669.
33. Jelsma,T.N., Howe,J.A., Evelegh,C.M., Cunniff,N.F., Skiadopoulos,M.H., Floroff,M.R., Denman,J.E. and Bayley,S.T. (1988) Use of deletion and point mutants spanning the coding region of the adenovirus 5 E1A gene to define a domain that is essential for transcriptional activation. *Virology*, **163**, 494–502.
34. Jelsma,T.N., Howe,J.A., Mymryk,J.S., Evelegh,C.M., Cunniff,N.F.A. and Bayler,S.T. (1989) Sequences in E1A proteins of human adenovirus 5 required for cell transfromation, repression of a transcriptional enhancer and induction of proliferating cell nuclear antigen. *Virology*, **171**, 120–130.
35. Howe,J.A., Mymryk,J.S., Egan,C., Branton,P.E. and Bayley,S.T. (1990) Retinoblastoma Growth Suppressor and a 300-kDa Protein Appear to Regulate Cellular DNA Synthesis. *Proc. Natl Acad. Sci. USA*, **87**, 5883–5887.
36. Harlow,E., Franza,B.R. and Schley,C. (1985) Monoclonal antibodies for adenovirus early region 1A proteins: extensive heterogeneity in early region 1A products. *J. Virol.*, **55**, 533–546.
37. Shih,M.-C., Heinrich,P. and Goodman,H.M. (1991) Cloning and chromosomal mapping of nuclear genes encoding chloroplast and cytosolic glyceraldehyde-3-phosphate-dehydrogenase from *Arabidopsis thaliana*. *Gene*, **104**, 133–138.
38. Arany,Z., Newsome,D., Oldread,E., Livingston,D.M. and Eckner,R. (1995) A family of transcriptional adaptor proteins targeted by the E1A oncoprotein. *Nature*, **374**, 81–84.
39. Yang,X.J., Ogryzko,V.V., Nishikawa,J., Howard,B.H. and Nakatani,Y. (1996) A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature*, **382**, 319–324.
40. Nakajima,T., Uchida,C., Anderson,S.F., Lee,C.G., Hurwitz,J., Parvin,J.D. and Montminy,M. (1996) The signal-dependent coactivator CBP is a nuclear target for pp90$^{RSK}$. *Cell*, **86**, 465–474.
41. Kraus,W.L., Manning,E.T. and Kadonaga,J.T. (1999) Biochemical analysis of distinct activation functions in p300 that enhance transcription initiation with chromatin templates. *Mol. Cell. Biol.*, **19**, 8123–8135.
42. Dhalluin,C., Carlson,J.E., Zeng,L., He,C., Aggarwal,A.K. and Zhou,M.M. (1999) Structure and ligand of a histone acetyltransferase bromodomain. *Nature*, **399**, 491–496.
43. Smith,E.R., Belote,J.M., Schiltz,R.L., Yang,X.-J., Moore,P.A., Berger,S.L., Nakatani,Y. and Allis,C.D. (1998) Cloning of *Drosophila* GCN5: conserved features among metazoan GCN5 family members. *Nucleic Acids Res.*, **26**, 2948–2954.
44. Ehrlich,K.C., Cary,J.W. and Ehrlich,M. (1992) A broad bean cDNA clone encoding a DNA-binding protein resembling mammalian CREB in its sequence specificity and DNA methylation sensitivity. *Gene*, **117**, 169–178.
45. Salvador,M.L., Sendra,R., Lòpez-Rodas,G., Tordera,V. and Franco,L. (1985) On the ubiquitous presence of histone acetyltransferase B in eukaryotes. *FEBS Lett.*, **191**, 55–58.
46. Lopez-Rodas,G., Georgieva,E.I., Sendra,R. and Loidl,P. (1991) Histone acetylation in Zea mays. I. Activities of histone acetyltransferases and histone deacetylases. *J. Biol. Chem.*, **266**, 18745–18750.
47. Georgieva,E.I., Lopez-Rodas,G., Sendra,R., Grobner,P. and Loidl,P. (1991) Histone acetylation in Zea mays. II. Biological significance of post-translational histone acetylation during embryo germination. *J. Biol. Chem.*, **266**, 18751–18760.
48. Grabher,A., Brosch,G., Sendra,R., Lechner,T., Eberharter,A., Georgieva,E.I., Lopez-Rodas,G., Franco,L., Dietrich,H. and Loidl,P. (1994) Subcellular location of enzymes involved in core histone acetylation. *Biochemistry*, **33**, 14887–14895.