Oncogene (2002) 21, 2253–2260

www.nature.com/onc

SHORT REPORT

Acetyltransferase machinery conserved in p300/CBP-family proteins

L Wuchao Yuan[1] and Antonio Giordano*[,2]

[1]Department of Physiology and Biophysics, Boston University School of Medicine, Boston, Massachusetts, MA 02118, USA; [2]Department of Pathology, Anatomy and Cell Biology, Jefferson Medical College, Philadelphia, Pennsylvania, PA 19107, USA

CREB-binding protein (CBP) and p300 are highly conserved and functionally related transcription coactivators and histone/protein acetyltransferases. They are tumor suppressors, participate in a wide variety of physiological events, and serve as integrators among different signal transduction pathways. In this study, 11 distinct proteins that have a high degree of homology with the amino acid sequence of p300 have been identified in current protein databases. All of these 11 proteins belong to either animal or plant multicellular organisms (higher eucaryotes). Conservation of p300/CBP domains among these proteins was examined further by sequence alignment and pattern search. The domains of p300/CBP that are required for the HAT function, including PHD, putative CoA-binding, and ZZ domains, are conserved in all of these 11 proteins. This observation is consistent with the previous functional assays and indicates that they are a family of acetyltransferases, i.e. *p*300/*C*BP *a*cetyl*t*ransferases (PCAT). TAZ domains (TAZ1 and/or TAZ2) of PCAT proteins may allow them to participate in transcription regulation by either directly recruiting transcription factors, acetylating them subsequently, or directing targeted acetylation of nucleosomal histones.
Oncogene (2002) **21,** 2253–2260. DOI: 10.1038/sj/onc/1205283

Keywords: p300; CBP; PCAT; histone/protein acetyltransferase; protein domains; protein structure

The cellular 300-kD protein (p300) was first discovered as one of the major cellular proteins that interact with adenoviral E1A transforming protein in 1986 (Harlow *et al*., 1986). The cDNA encoding human p300 protein was cloned in 1994 (Eckner *et al*., 1994b). The predicted amino acid sequence of p300 has a high degree of homology with those of CREB-binding protein (CBP) (Chrivia *et al*., 1993) and another hypothetical protein CBP-1 in *Caenorhabditis elegans* database (Shi and Mello, 1998; Wilson *et al*., 1994). Thus, they were proposed to belong to the same protein family (Arany *et al*., 1994). The fact that both p300 and CBP bound to E1A and coactivated CREB-mediated transcription supports this proposal. Later, this protein family was expanded after the discovery of *Drosophila* CBP (dCBP), which was also able to bind to E1A and coactivate CREB-dependent transactivation (Akimaru *et al*., 1997a). Recently, p300/CBP-like proteins of the plant *Arabidopsis thaliana* have been reported (Bordoli *et al*., 2001), suggesting that they are possible members of the p300/CBP family. It is interesting to know how large the p300/CBP protein family could be and how this family of proteins is conserved evolutionarily.

Taking advantage of recent progress in genomic sequencing of different organisms, we performed a blast search (Altschul *et al*., 1990) with the WU-Blast 2.0 software (Yuan *et al*., 1998) using human p300 protein sequence as bait against current protein databases covering different organisms. Eleven distinct proteins homologous to p300 were recovered, with *P*-values less than 10^{-23}. The *P*-values for the other proteins recovered in the search were higher than 10^{-2}. Therefore, the 11 homologous proteins are proposed to be part of the same family (Table 1). They are either animal or plant proteins. Besides five of them that have been cloned and are animal proteins, five others were found in *Arabidopsis thaliana*, and the other one in *Oryzae sativas* (rice). In addition, a 348-AA fragment of *Zea mays* that has significant homology with the HAT domain sequence of p300/CBP was also found in Plant Chromatin Database (ChromDB, http://Ag.Arizona.edu/chromatin/chromatin.html). Interestingly, no p300/CBP-family protein was found in single-cell organisms, suggesting that p300/CBP function is associated with physiological pathways defining multicellular organisms during development. Consistent with this, the mice lacking either p300 or CBP and the fruit flies defective in dCBP lead to embryonic lethality, dying at the early stage (Akimaru *et al*., 1997b; Yao *et al*., 1998; Giordano and Avantaggiati, 1999; Puri and Sartorelli, 2000).

p300 and CBP were initially identified as transcriptional coactivators (Eckner *et al*., 1994a; Kwok *et al*., 1994; Lundblad *et al*., 1995), which were defined as adapter or integrator between DNA-binding activators and the basal transcriptional machinery (Goodrich and Tjian, 1994). In support of this, they have been found to interact with a variety of transcriptional activators (Goodman and Smolik, 2000; Shikama *et al*., 1997), as well as components of the basal transcriptional machinery, such as the TATA-binding protein (TBP), TFIIB, and RNA polymerase II (RNAPII) (Abraham *et al*., 1993; Kwok *et al*., 1994; Nakajima *et al*., 1997; Yuan *et*

*Correspondence: A Giordano; E-mail: agiordan@lac.jci.tju.edu
Received 6 September 2001; revised 18 December 2001; accepted 3 January 2002

Table 1 p300/CBP acetyltransferase (PCAT) family of proteins*

Organism	*Protein name*	*Other name*	*Residues*	*Accession#*
Homo sapiens (Human)	p300 (Eckner *et al*., 1994b)	E1A-binding protein	2414 AA	Q09472
	hCBP(Giles *et al*., 1997)	CREB-binding protein	2442 AA	AAC51331
Mus musculus (Mouse)	mCBP (Chrivia *et al*., 1993)	CREB-binding protein	2441 AA	P45481
Drosophila melanogaster (Fruit fly)	dCBP (Akimaru *et al*., 1997a)	CG15319 (Adams *et al*., 2000)	3275 AA	Q9W321
Caenorhabditis elegans	CBP-1 (Wilson *et al*., 1994)	R10E11.1	2056 AA	P34545
	HAC1	PCAT2 (Bordoli *et al*., 2001)	1654 AA	AAG60059, HAC000001 (ChromDB)
	HAC2	PCAT1 (Bordoli *et al*., 2001) F1N21.18	1689 AA	HAC000002 (ChromDB)
Arabidopsis thaliana	HAC4	PCAT3 (Bordoli *et al*., 2001) F14J16.27	1550 AA	Q9LG11, HAC000004 (ChromDB)
	HAC5	PCAT4 (Bordoli *et al*., 2001)	1670 AA	Q9LE42, HAC000005 (ChromDB)
	1498-AA protein	PCAT5, F17F16.8	1498 AA	Q9FWQ5
Oryza sativa (Rice)	rCBP	Putative CBP-related acetyltransferase	994 AA	Q9XHY7
Zea mays	zCBP	Putative CBP-related acetyltransferase	348 AA (partial)	HAC000101 (ChromDB)

*The proteins homologues to p300 were searched by using the Washington University BLAST software 2.0 version [Rel. 2.0MP-WashU, 09-Sep-1999] (advanced WU-blast search) at Bork's group in EMBL (Heidelberg) (Yuan *et al*., 1998). In the search, the amino acid sequence of human p300 protein was used as a bait. The protein databases used in the search were provided by the ExPASy (Expert Protein Analysis System) proteomics server of the Swiss Institute of Bioinformatics (SIB), including nrdb95, swissprot, pdb, sp-nrdb and nrdb. The similarity of the recovered proteins to p300 was determined by their *P*-values provided by the search. Some homology information used in this study was obtained from the Plant Chromatin Database

al., 1996). Because the wide variety of transcriptional activators important to many different signal transduction pathways use p300 and CBP as coactivators, the cellular demand for p300/CBP is so high that they may be used competitively among these pathways and are limiting factors at least in some of the pathways. Therefore, p300/CBP plays the role of a real cellular integrator and regulator in physiological processes (Akimaru *et al*., 1997b; Yao *et al*., 1998).

The breakthrough in understanding p300/CBP function resulted from identification of their intrinsic histone acetyltransferase (HAT) activities (Bannister and Kouzarides, 1996; Ogryzko *et al*., 1996). They not only use nucleosomal histones as substrates but also many other transcription factors (Boyes *et al*., 1998; Deng *et al*., 2000; Gu and Roeder, 1997; Imhof *et al*., 1997; Naar *et al*., 1998; Zhang and Bieker, 1998; Cress and Seto, 2000). p300/CBP can also acetylate itself (Kraus and Kadonaga, 1998). In most of these cases, acetylation has been found to affect p300/CBP function and its participated transcription.

In comparison with the intensive studies of p300/CBP HAT and transcriptional coactivator activities, studies of the structural basis of their function have been relatively few and preliminary. p300 and CBP contain multiple conserved domains, such as the CREB-binding domain (KIX) (Lundblad *et al*., 1995; Radhakrishnan *et al*., 1999), bromodomain, and three cysteine/histidine-rich regions (CH1, CH2 and CH3) (Eckner *et al*., 1994a). Three CH regions are composed of four zinc finger motifs which are TAZ1 in CH1 (Ponting *et al*., 1996), PHD in CH2 (Aasland *et al*., 1995; Koken *et al*., 1995), and ZZ and TAZ2 in CH3 (Ponting *et al*., 1996). Besides, several other conserved domains have also been found in mammalian p300 and CBP. An N-terminal nuclear receptor-interacting domain (RID) interacts with nuclear receptor activators (ER, RAR, RXR, T3R) (Chakravarti *et al*., 1996; Kamei *et al*., 1996). A bipartite nuclear localization signal (NLS-BP) resides at the middle part of p300/CBP (Yuan and Gambee, 2000). A glutamine/proline-rich (QP) domain near the C-terminus associates with other coactivator and HAT proteins P/CIP and SRC-1 (Fontes *et al*., 1999). A C-terminal proline-rich (P) domain is also conserved but there are no reports about its function yet.

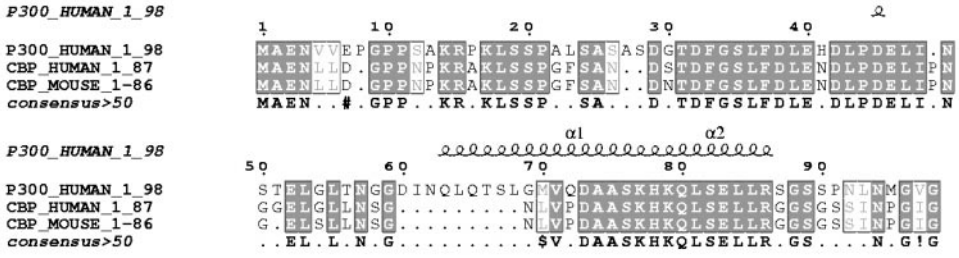
The function of p300/CBP is executed through different domains. RID, KIX, TAZ1 and TAZ2 bind to different sets of transcriptional activators and regulators (Goodman and Smolik, 2000). These interactions are required for p300/CBP to be recruited onto DNA. These domains are essential to p300/CBP targeting acetylation of nucleosomes on chromatin templates and transcription coactivator function (Kundu *et al*., 2000; Utley *et al*., 1998). Both N- and C-terminal regions of p300/CBP (e.g., 1–569 and 1737–2414 of human p300) contain transcription activation (TA) activity when they are fused to the DNA binding domain (Chrivia *et al*., 1993; Yuan *et al*., 1996). Although no defined domains have been experimentally confirmed to be responsible for TA function, TAZ2 binding to TFIIB and RNA helicase A (Kwok *et al*., 1994; Yuan *et al*., 1996) and QP domain binding to P/CIP and SRC-1 (Fontes *et al*., 1999) have the potential to participate in the intrinsic TA activity of p300/CBP.

The HAT function of p300 and CBP is associated with a large conserved region, spanning from the PHD domain to the ZZ motif (Bannister and Kouzarides, 1996; Ogryzko *et al*., 1996). In the middle of the HAT domain there is a conserved region with certain homology to the coenzyme A (CoA)-binding domain of PCAF and GCN5 (Clements *et al*., 1999; Rojas *et al*., 1999). The function of the putative

CoA-binding domain in p300/CBP has not been confirmed yet, and it is still unknown how the PHD and ZZ domains are involved in the HAT function. In PCAF, the bromodomain has been shown to bind to acetylated lysine (Dhalluin *et al*., 1999), implying that it may be involved in HAT function. However, the bromodomain of p300/CBP is not required for the HAT function (Bannister and Kouzarides, 1996; Ogryzko *et al*., 1996).

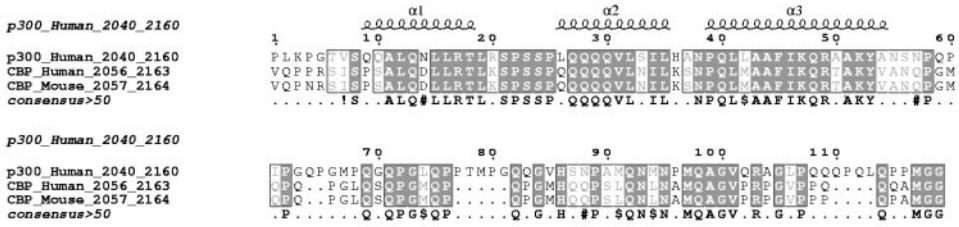
To find out how p300/CBP domains are conserved among 11 proteins of the p300/CBP family, we aligned these protein sequences by different groups with CLUSTALW (Thompson *et al*., 1994), and the alignments were tuned up with UCSC SAM-T99 protocol (Karplus *et al*., 1998). We also conducted a pattern search with PATTINPROT protocol on NPS@ server (Combet *et al*., 2000) using the pattern sequences of each domain or motif against these 11 protein sequences. Both methods yielded similar and consistent results. The p300/CBP domains that were conserved only in mammalian systems included the RID, NLS – BP, and the C-terminal QP and P domains (Figure 1a). The KIX domain and bromodomain were shared only by all animal proteins of the p300/CBP family (Figure 1b). The TAZ domains (TAZ1 and TAZ2) that interacted with a variety of transcriptional activators are conserved among almost all of the p300/CBP-family proteins with only two exceptions, Arabidopsis HAC1 and the rice rCBP (Figure 1c). The domains required for the HAT function including the PHD, putative CoA-binding, and ZZ zinc finger domains were conserved absolutely in all proteins of the family (Figure 1d). Many of those in *Arabidopsis* contained two copies of the ZZ motif. In addition, the HAT activity of *Arabidopsis* HAC1 and HAC2 has been confirmed experimentally (Bordoli *et al*., 2001). Therefore, it is reasonable to refer to

a **Receptor-interacting domain (RID)**

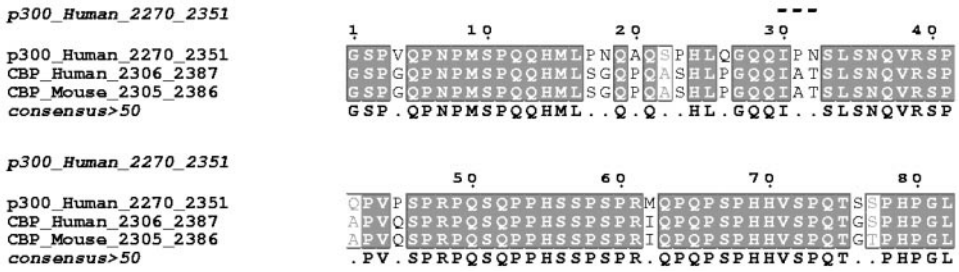


Bipartite nuclear localization signal (NLS-BP)

Glutamine- and proline-rich domain (QP)



C-terminal proline-rich domain (P)



(*Continued*)

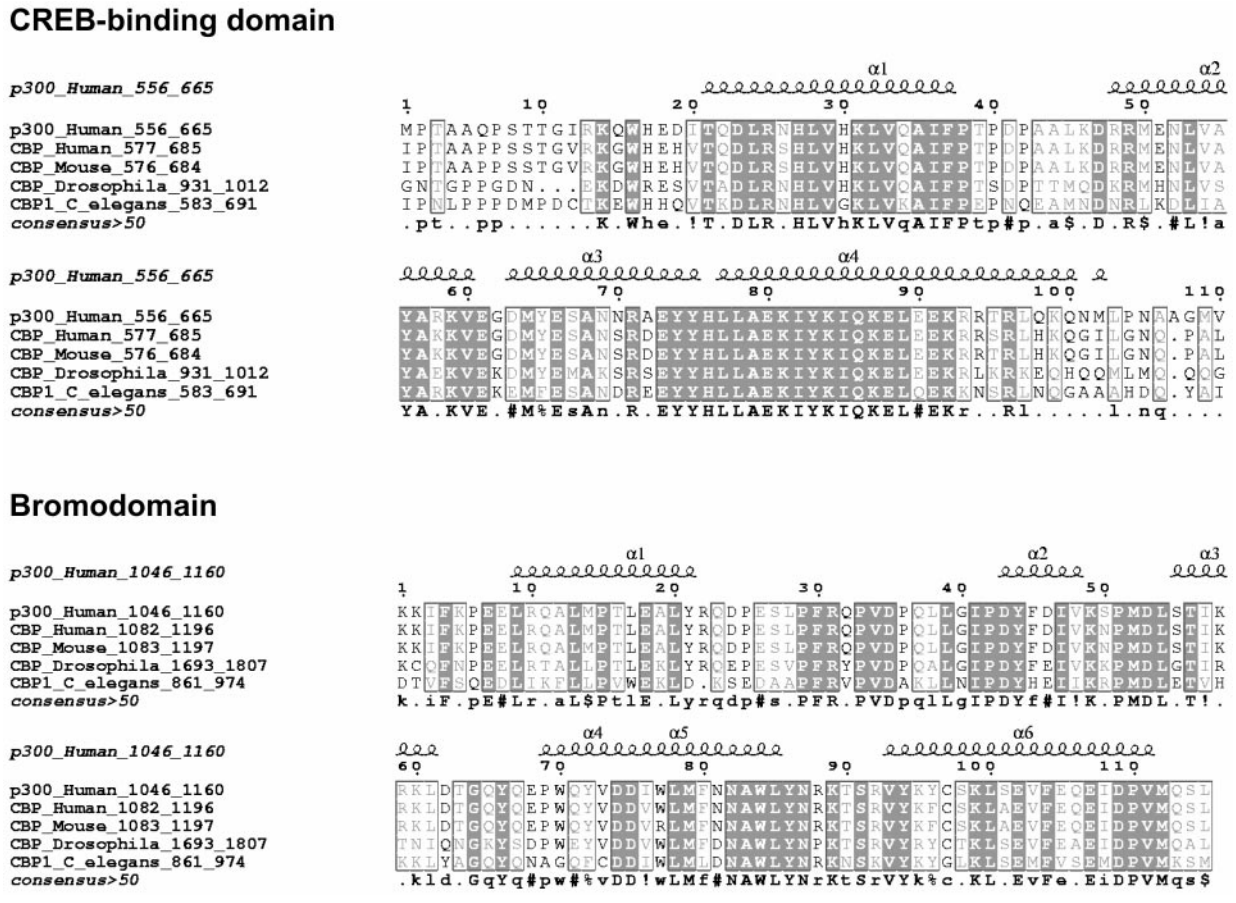
these p300/CBP family proteins as p300/CBP acetyltransferases (PCAT) (Bordoli *et al*., 2001). It still needs to be tested whether they also use other transcription factors as substrates during acetylation.

Although the putative CoA-binding domain of PCAT has certain homology with that of GNAT (GCN5-related N-acetyltransferase) proteins including GCN5 and PCAF (Neuwald and Landsman, 1997) (Figure 1e), the PCAT CoA-binding domain is highly conserved within PCAT proteins (Figure 1d). The predicted secondary structures between these two CoA-binding domains are also distinct (Clements *et al*., 1999). These observations suggest that the PCAT CoA-binding domain belongs to a separate family. So far, the 3-D structures of the PCAT-related domains, such as the PHD, CoA-binding and ZZ domains, and the mechanisms for them to function in acetylation have not yet been reported.

PCAT proteins can acetylate histones and other transcription factors. Recruitment of PCAT proteins onto chromatin by DNA-binding activators is required for targeting acetylation of nucleosomal histones (Kundu *et al*., 2000; Utley *et al*., 1998). It is possible for PCAT proteins to acetylate other transcription factors only if they interact with each other. Thus, the conserved TAZ domains would make the acetylation potential of PCAT proteins functionable. However, not all of PCAT proteins contain TAZ domains, suggesting that they may have other mechanisms to access substrates.

The ability of PCAT proteins to acetylate histones and other transcription factors allow them to remodel chromatin and modulate activity of the transcription factors. These functions are essential for p300 and CBP to serve as transcription coactivators. Because PCAT proteins have acetylation function, they are likely to have transcription coactivator activity too. This hypothesis needs to be tested. Besides the acetylation-dependent mechanisms in transcription coactivation, mammalian p300 and CBP also contain intrinsic transcriptional activity (Chrivia *et al*., 1993; Yuan *et al*., 1996). The remaining question is whether other PCAT proteins contain similar intrinsic transcriptional activity when fused to a DNA-binding domain and participate in transcription regulation through a HAT-independent mechanism.

In summary, p300- and CBP-related proteins are all found in multi-cellular organisms and belong to a distinct family. Because of the conserved acetylation-related domains in all members of this protein family and other experimental evidence, they are believed to be acetyltransferases and therefore are called as PCAT. The potential of PCAT proteins to acetylate histones and other transcription factors suggests that they are transcription coactivators as well. Because conserved

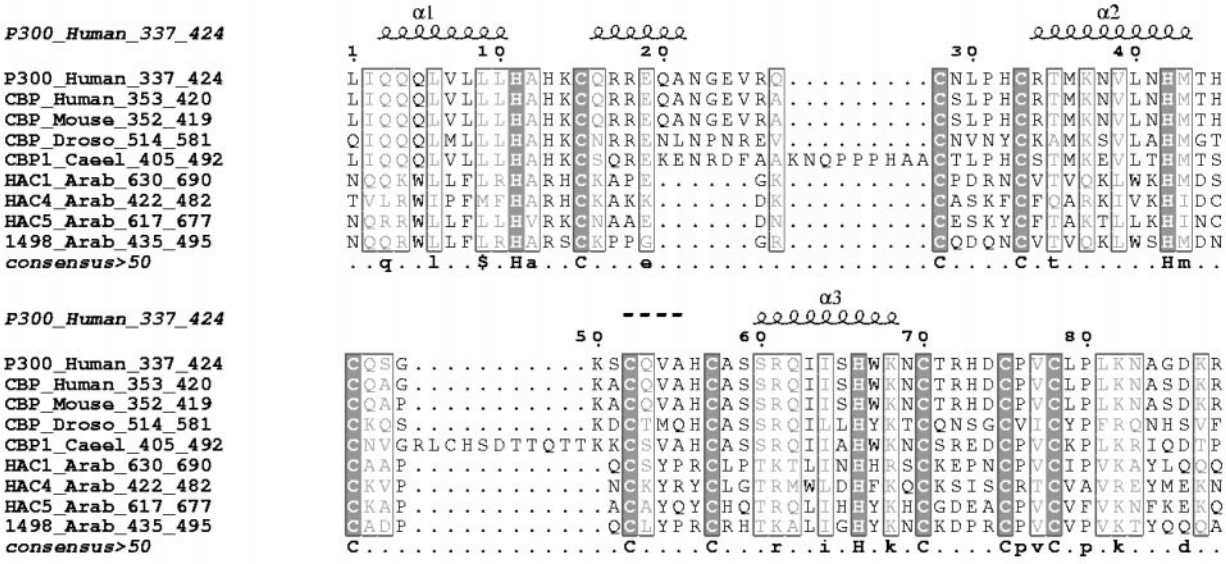


(*Continued*)

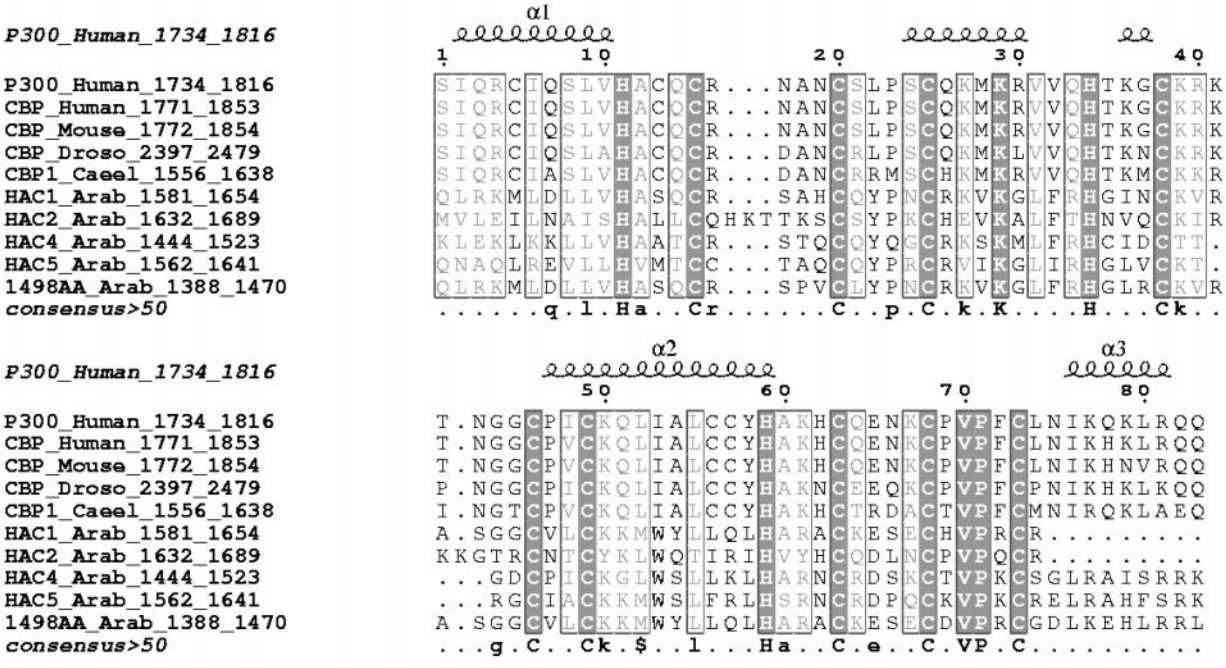
TAZ domains of PCAT proteins interact with a variety of transcription activators, the PCAT proteins are expected to be integrators to regulate multiple signal transduction pathways during development.

Acknowledgments
We thank Ms Marie L Basso for her editorial assistance during manuscript preparation. This work was supported by NIH grants to A Giordano.

c **Transcriptional adaptor zinc finger 1 (TAZ1)**

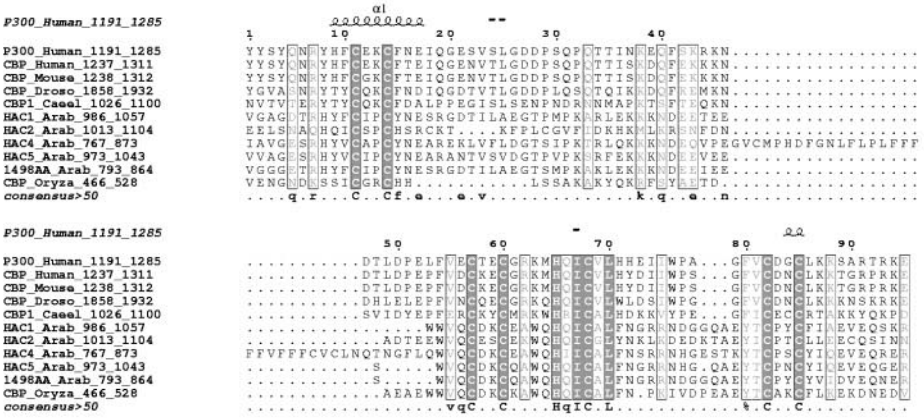


Transcriptional adaptor zinc finger 2 (TAZ2)

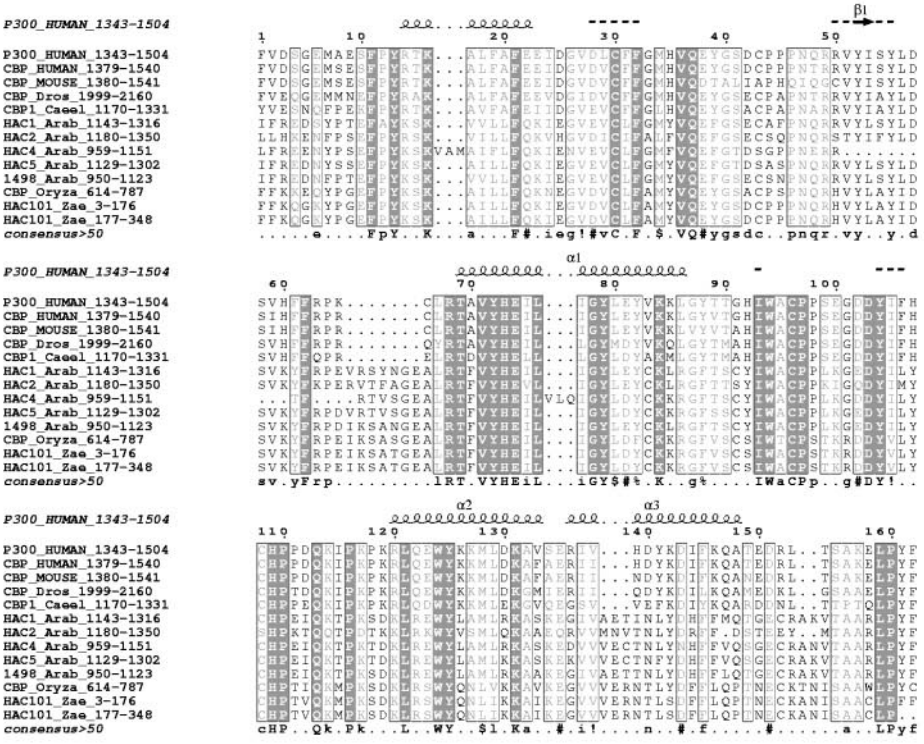


(*Continued*)

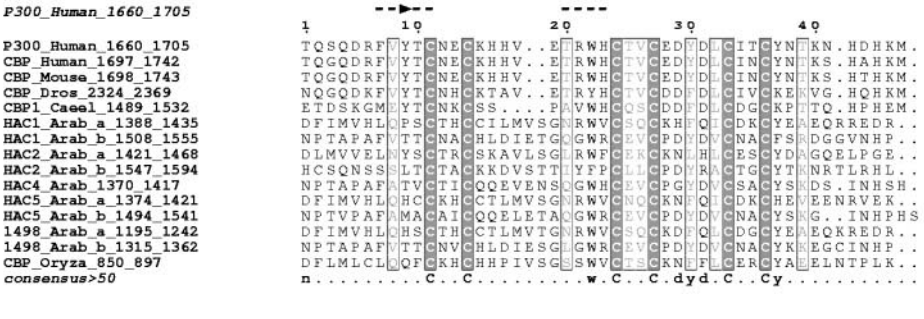
d Plant homeodomain (PHD) motif



Putative CoA-binding domain



ZZ zinc finger motif



(*Continued*)

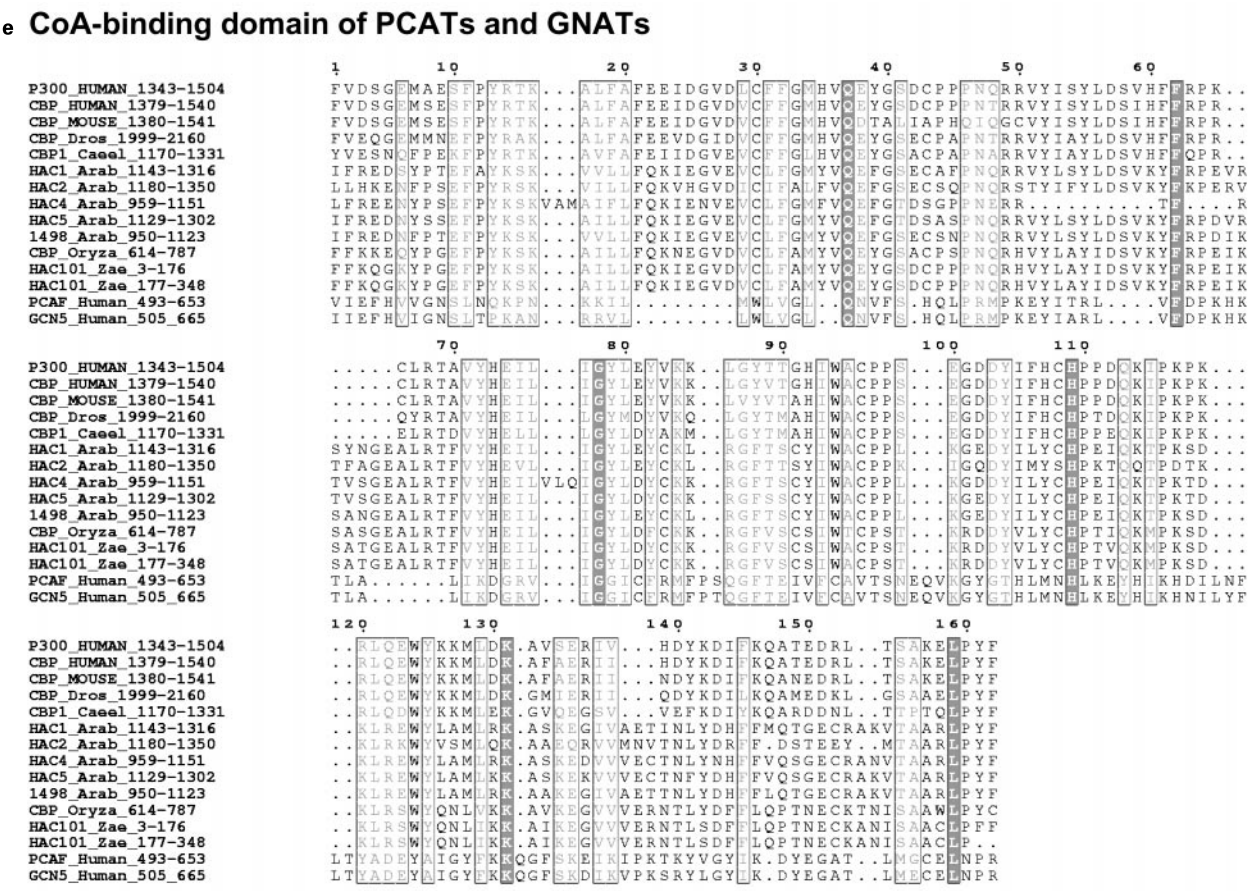


Figure 1 The multiple protein sequence alignment of the p300/CBP-family proteins were carried out by using the CLUSTALW protocol (Thompson *et al*., 1994). In the alignment, the p300/CBP-family proteins were gathered by different groups at different levels of the taxonomic relation. The alignments were tuned up further by using the Sequence Alignment and Modeling System (UCSC SAM-T99 program) provided by the Computational Biology Group at University of California-Santa Cruz (Karplus *et al*., 1998). The protein pattern search was performed by using the PATTINPROT protocol on Network Protein Sequence Analysis (NPS@) server at the PBIL Lyon-Gerland (Combet *et al*., 2000). In the pattern search, the sequences of the known p300/CBP domains and motifs were used against all of the p300/CBP-family proteins that we recognized in this study. The secondary structure of these proteins were analysed by the Profile Network Prediction HeiDelberg (PHD) (Rost and Sander, 1993) using PredictProtein service from Columbia University. The sequence alignment is displayed with the ESPript 1.9 program. The top line of each alignment block is the secondary structure predicted by PHD protocol (Rost and Sander, 1993). Helix symbols indicate α-helices. Dashed line with or without arrows indicates β-sheet. The labels 'α' and 'β' indicate the predicted structures with high probability. The bottom line of each alignment block represents the consensus sequence generated using MULTALIN Algorithm (Corpet, 1988). Uppercase is identity. Lowercase is consensus level >0.5. '!' is I or V. '$' is L or M. '%' is F or Y. '#' is either one of NDQEBZ. (**a**) The domains conserved only in human and mouse p300 and CBP. (**b**) The domains of the p300/CBP-family proteins conserved within animal systems. (**c**) The TAZ zinc finger domains that are conserved among almost all of the p300/CBP-family proteins. (**d**) The domains conserved in all of the p300/CBP-family proteins. (**e**) Amino acid sequence alignment of the CoA-binding domains of PCAT proteins and GCN5 and PCAF

References

Aasland R, Gibson TJ and Stewart AF. (1995). *Trends Biochem. Sci.,* **20,** 56 – 59.

Abraham SE, Lobo S, Yaciuk P, Wang HG and Moran E. (1993). *Oncogene,* **8,** 1639 – 1647.

Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, George RA, Lewis SE, Richards S, Ashburner M, Henderson SN, Sutton GG, Wortman JR, Yandell MD, Zhang Q, Chen LX, Brandon RC, Rogers YH, Blazej RG, Champe M, Pfeiffer BD, Wan KH, Doyle C, Baxter EG, Helt G, Nelson CR, Gabor GL, Abril JF, Agbayani A, An HJ, Andrews-Pfannkoch C, Baldwin D, Ballew RM, Basu A, Baxendale J, Bayraktaroglu L, Beasley EM, Beeson KY, Benos PV, Berman BP, Bhandari D, Bolshakov S, Borkova D, Botchan MR, Bouck J, Brokstein P, Brottier P, Burtis KC, Busam DA, Butler H, Cadieu E, Center A, Chandra I, Cherry JM, Cawley S, Dahlke C, Davenport LB, Davies P, de Pablos B, Delcher A, Deng Z, Mays AD, Dew I, Dietz SM, Dodson K, Doup LE, Downes M, Dugan-Rocha S, Dunkov BC, Dunn P, Durbin KJ, Evangelista CC, Ferraz C, Ferriera S, Fleischmann W, Fosler C, Gabrielian AE, Garg NS, Gelbart WM, Glasser K, Glodek A, Gong F, Gorrell JH, Gu Z, Guan P, Harris M, Harris NL, Harvey D, Helman TJ, Hernandez JR, Houck J, Hostin D, Houston KA, Howland TJ, Wei MH, Ibegwam C, Jalali M, Kalush F, Karpen GH, Ke Z, Kennison JA, Ketchum KA, Kimmel BE, Kodira CD, Kraft C, Kravitz S, Kulp D, Lai

Z, Lasko P, Lei Y, Levitsky AA, Li J, Li Z, Liang Y, Lin X, Liu X, Mattei B, McIntosh TC, McLeod MP, McPherson D, Merkulov G, Milshina NV, Mobarry C, Morris J, Moshrefi A, Mount SM, Moy M, Murphy B, Murphy L, Muzny DM, Nelson DL, Nelson DR, Nelson KA, Nixon K, Nusskern DR, Pacleb JM, Palazzolo M, Pittman GS, Pau S, Pollard J, Puri V, Reese MG, Reinert K, Remington K, Saunders RD, Scheeler F, Shen H, Shue BC, Siden-Kiamos I, Simpson M, Skupski MP, Smith T, Spier E, Spradling AC, Stapleton M, Strong R, Sun E, Svirskas R, Tector C, Turner R, Venter E, Wang AH, Wang X, Wang ZY, Wassarman DA, Weinstock GM, Weissenbach J, Williams SM, Woodage T, Worley KC, Wu D, Yang S, Yao QA, Ye J, Yeh RF, Zaveri JS, Zhan M, Zhang G, Zhao Q, Zheng L, Zheng XH, Zhong FN, Zhong W, Zhou X, Zhu S, Zhu X, Smith HO, Gibbs RA, Myers EW, Rubin GM and Venter JC (2000). *Science,* **287,** 2185–2195.

Akimaru H, Chen Y, Dai P, Hou DX, Nonaka M, Smolik SM, Armstrong S, Goodman RH and Ishii S. (1997a). *Nature,* **386,** 735–738.

Akimaru H, Hou DX and Ishii S. (1997b). *Nat. Genet.,* **17,** 211–214.

Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ. (1990). *J. Mol. Biol.,* **215,** 403–410.

Arany Z, Sellers WR, Livingston DM and Eckner R. (1994). *Cell,* **77,** 799–800.

Bannister AJ and Kouzarides T. (1996). *Nature,* **384,** 641–643.

Bordoli L, Netsch M, Luthi U, Lutz W and Eckner R. (2001). *Nucleic Acids Res.,* **29,** 589–597.

Boyes J, Byfield P, Nakatani Y and Ogryzko V. (1998). *Nature,* **396,** 594–598.

Chakravarti D, LaMorte VJ, Nelson MC, Nakajima T, Schulman IG, Juguilon H, Montminy M and Evans RM. (1996). *Nature,* **383,** 99–103.

Chrivia JC, Kwok RP, Lamb N, Hagiwara M, Montminy MR and Goodman RH. (1993). *Nature,* **365,** 855–859.

Clements A, Rojas JR, Trievel RC, Wang L, Berger SL and Marmorstein R. (1999). *EMBO J.,* **18,** 3521–3532.

Combet C, Blanchet C, Geourjon C and Deleage G. (2000). *Trends Biochem. Sci.,* **25,** 147–150.

Corpet F. (1988). *Nucleic Acids Res.,* **16,** 10881–10890.

Cress WD and Seto E. (2000). *J. Cell. Physiol.,* **184,** 1–16.

Deng L, de la Fuente C, Fu P, Wang L, Donnelly R, Wade JD, Lambert P, Li H, Lee CG and Kashanchi F. (2000). *Virology,* **277,** 278–295.

Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK and Zhou MM. (1999). *Nature,* **399,** 491–496.

Eckner R, Arany Z, Ewen M, Sellers W and Livingston DM. (1994a). *Cold Spring Harb. Symp. Quant. Biol.,* **59,** 85–95.

Eckner R, Ewen ME, Newsome D, Gerdes M, DeCaprio JA, Lawrence JB and Livingston DM. (1994b). *Genes Dev.,* **8,** 869–884.

Fontes JD, Kanazawa S, Jean D and Peterlin BM. (1999). *Mol. Cell. Biol.,* **19,** 941–947.

Giles RH, Petrij F, Dauwerse HG, den Hollander AI, Lushnikova T, van Ommen GJ, Goodman RH, Deaven LL, Doggett NA, Peters DJ and Breuning MH. (1997). *Genomics,* **42,** 96–114.

Giordano A and Avantaggiati ML. (1999). *J. Cell. Physiol.,* **181,** 218–230.

Goodman RH and Smolik S. (2000). *Genes Dev.,* **14,** 1553–1577.

Goodrich JA and Tjian R. (1994). *Curr. Opin. Cell. Biol.,* **6,** 403–409.

Gu W and Roeder RG. (1997). *Cell,* **90,** 595–606.

Harlow E, Whyte P, Franza BR and Schley C. (1986). *Mol. Cell. Biol.,* **6,** 1579–1589.

Imhof A, Yang XJ, Ogryzko VV, Nakatani Y, Wolffe AP and Ge H. (1997). *Curr. Biol.,* **7,** 689–692.

Kamei Y, Xu L, Heinzel T, Torchia J, Kurokawa R, Gloss B, Lin SC, Heyman RA, Rose DW, Glass CK and Rosenfeld MG. (1996). *Cell,* **85,** 403–414.

Karplus K, Barrett C and Hughey R. (1998). *Bioinformatics,* **14,** 846–856.

Koken MH, Saib A and de The H. (1995). *C R Acad. Sci. III,* **318,** 733–739.

Kraus WL and Kadonaga JT. (1998). *Genes Dev.,* **12,** 331–342.

Kundu TK, Palhan VB, Wang Z, An W, Cole PA and Roeder RG. (2000). *Mol. Cell,* **6,** 551–561.

Kwok RP, Lundblad JR, Chrivia JC, Richards JP, Bachinger HP, Brennan RG, Roberts SG, Green MR and Goodman RH. (1994). *Nature,* **370,** 223–226.

Lundblad JR, Kwok RP, Laurance ME, Harter ML and Goodman RH. (1995). *Nature,* **374,** 85–88.

Naar AM, Beaurang PA, Robinson KM, Oliner JD, Avizonis D, Scheek S, Zwicker J, Kadonaga JT and Tjian R. (1998). *Genes Dev.,* **12,** 3020–3031.

Nakajima T, Uchida C, Anderson SF, Lee CG, Hurwitz J, Parvin JD and Montminy M. (1997). *Cell,* **90,** 1107–1112.

Neuwald AF and Landsman D. (1997). *Trends Biochem. Sci.,* **22,** 154–155.

Ogryzko VV, Schiltz RL, Russanova V, Howard BH and Nakatani Y. (1996). *Cell,* **87,** 953–959.

Ponting CP, Blake DJ, Davies KE, Kendrick-Jones J and Winder SJ. (1996). *Trends Biochem. Sci.,* **21,** 11–13.

Puri PL and Sartorelli V. (2000). *J. Cell Physiol.,* **185,** 155–173.

Radhakrishnan I, Perez-Alvarado GC, Parker D, Dyson HJ, Montminy MR and Wright PE. (1999). *J. Mol. Biol.,* **287,** 859–865.

Rojas JR, Trievel RC, Zhou J, Mo Y, Li X, Berger SL, Allis CD and Marmorstein R. (1999). *Nature,* **401,** 93–98.

Rost B and Sander C. (1993). *J. Mol. Biol.,* **232,** 584–599.

Shi Y and Mello C. (1998). *Genes Dev.,* **12,** 943–955.

Shikama N, Lyon J and La Thangue NB. (1997). *Trends Cell Biol.,* **7,** 230–236.

Thompson JD, Higgins DG and Gibson TJ. (1994). *Nucleic Acids Res.,* **22,** 4673–4680.

Utley RT, Ikeda K, Grant PA, Cote J, Steger DJ, Eberharter A, John S and Workman JL. (1998). *Nature,* **394,** 498–502.

Wilson R, Ainscough R, Anderson K, Baynes C, Berks M, Bonfield J, Burton J, Connell M, Copsey T, Cooper J, Coulson A, Craxton H, Dear S, Du Z, Durbin R, Favello A, Frazer A, Fulton L, Gardner A, Green P, Hawkins T, Hillier L, Jior M, Johnston L, Jones M, Kershaw J, Kirsten J, Laisster N, Latreille P, Lightning J, Lloyd C, Mortimer B, O'Callaghan M, Parsons J, Percy C, Rifken L, Roopra A, Saunders D, Shownkeen R, Simms M, Smaldon N, Smith A, Smith M, Sonnhammer E, Staden R, Sulston J, Thierry-Mieg J, Thomas K, Vaudin M, Vaughan K, Waterston R, Watson A, Weinstock L, Wilkinson-Sproat J and Wohldman P. (1994). *Nature,* **368,** 32–38.

Yao TP, Oh SP, Fuchs M, Zhou ND, Ch'ng LE, Newsome D, Bronson RT, Li E, Livingston DM and Eckner R. (1998). *Cell,* **93,** 361–372.

Yuan LW and Gambee JE. (2000). *J. Biol. Chem.,* **275,** 40946–40951.

Yuan W, Condorelli G, Caruso M, Felsani A and Giordano A. (1996). *J. Biol. Chem.,* **271,** 9009–9013.

Yuan YP, Eulenstein O, Vingron M and Bork P. (1998). *Bioinformatics,* **14,** 285–289.

Zhang W and Bieker JJ. (1998). *Proc. Natl. Acad. Sci. USA,* **95,** 9855–9860.