CHAPTER 1

CtBP Family Proteins:
Unique Transcriptional Regulators in the Nucleus with Diverse Cytosolic Functions

G. Chinnadurai*

Abstract

CtBP family proteins are unique in animals and in plants. The invertebrates and plants contain a single CtBP family gene while vertebrates have two genes. Genetic studies in *Drosophila* and in mice indicate that CtBPs play pivotal roles in animal development. The vertebrate CtBPs (CtBP1 and CtBP2) are highly related and are functionally redundant for certain developmental processes and non redundant for others. The vertebrates code two isoforms of each CtBP1 and CtBP2. The animal CtBPs exhibit a highly conserved sequence and structural similarity to D-isomer specific 2-hydroxy acid dehydrogenases (D2-HDH). Structural and molecular modeling studies indicate that CtBP1 is a dehydrogenase and could also bind with acyl-CoA under a different configuration. The CtBP family members function predominantly as transcriptional corepressors in the nucleus in conjunction with a number of different DNA binding repressors. The transcriptional regulatory activity of CtBPs appears to be regulated by NAD(H)-binding and the metabolic status of the cell. The corepressor complex of CtBP1 contains enzymatic constituents that mediate coordinated histone modification by deacetylation and methylation of histone H3-K9 and demethylation of histone H3-K4. In the cytosol, they perform diverse functions associated with membrane trafficking, central nervous system synapses and in regulation of the microtubule cytoskeleton. The mammalian CtBPs modulate oncogenesis by regulating the activities of tumor suppressor genes and cellular and viral oncogenes, consistent with a role in tumor suppression as well as in tumor promotion. The CtBPs promote tumorigenesis by repressing transcription of several critical pro-apoptotic genes and by inhibiting genes involved in the regulation of epithelial to mesenchymal transition. This Chapter presents a comprehensive general review of the CtBP field and highlights contents of the individual Chapters of this book which contain detailed discussions on structure and functions of animal and plant CtBP family proteins.

Introduction

CtBP (C-terminal binding protein) was identified in 1993 as a 48 kD cellular phosphoprotein that bound to the C-terminal region of the adenovirus E1A oncoprotein.[1] In 1995, the cDNA for the founding member of the CtBP family protein was cloned and the encoded protein was shown to bind to a five amino acid motif (PLDLS) conserved at the C-terminus of E1A of all primate adenoviruses.[2] The CtBP protein originally identified as the E1A-binding protein is now known as CtBP1. Subsequently, a highly homologous human protein termed CtBP2 was

*G. Chinnadurai—Institute for Molecular Virology. Saint Louis University Health Sciences Center. 3681 Park Avenue, St. Louis, Missouri 63110, U.S.A. Email: chinnag@slu.edu

CtBP Family Proteins, edited by G. Chinnadurai.

identified by analysis of EST data bank sequences[3] and mouse CtBP2 was cloned by a two hybrid screen against the transcription factor BKLF.[4] The initial amino acid homology searches revealed that CtBP1 shared a striking homology to D-isomer specific 2-hydroxy acid dehydrogenases (D2-HDH).[2] The interaction between a cellular protein with a metabolic enzyme fold and the E1A viral oncoprotein was unexpected since E1A functions primarily as a transcriptional modulator (reviewed by Gallimore and Turnell).[5] However, a possible role of CtBP in transcriptional repression was soon suggested by a tethering transcriptional assay.[6] In these assays, the N-terminal conserved region (CR1) of E1A fused to a heterologous DNA-binding domain (Gal4) strongly activated a synthetic promoter containing a Gal4 binding site. The CR1 region of E1A contains the sequences for interaction with a SWI/SNF-related chromatin remodeling complex, TRRAP/p400[7-9] and also the binding sites for the nuclear acetylase P/CAF.[10] Inclusion of the C-terminal region of E1A in the chimeric Gal4-E1A construct abrogated CR1-mediated transcriptional activation. Deletion of the CtBP-binding motif relieved the repressive activity of the C-terminal region. These results suggested that interaction of CtBP with the C-terminal region antagonized the trans-activation activity of CR1 in *cis*.

A definitive role for CtBP in transcriptional repression became evident with the identification and cloning of the *Drosophila* homolog of CtBP (dCtBP) by the laboratories of Michael Levine[11] and Susan Parkhurst.[12] Since then, a large number of DNA-binding transcriptional repressors have been reported to recruit CtBPs via the PLDLS-related binding sites.[13,14] The studies with dCtBP and a number of subsequent studies with vertebrate CtBP1 and CtBP2 have established that CtBPs function predominantly as transcriptional corepressors. However, splice variants of the vertebrate CtBPs have been shown to be involved in unrelated biological processes in the cytosol. During the past ten years since the cloning of CtBP1, there has been a substantial increase in our understanding of the structure, functions, and mechanisms of action of CtBP family proteins and their role in various biological processes. These advancements include elucidation of the structural determinants of CtBP1 and the molecular basis of its interaction with the CtBP-binding motif and the determination of the roles of CtBP1 and CtBP2 in mouse development. Additionally, several nuclear cofactors that mediate the transcriptional regulatory activity of CtBPs and the CtBP-target genes have been identified. This Chapter will highlight the salient aspects of CtBP family proteins while more detailed discussions can be found in the individual Chapters of this book.

CtBP Family Proteins

The CtBP family proteins are highly conserved in higher eukaryotes. The genomes of invertebrates such as *Drosophila* and *C. elegans* contain a single *CtBP* gene. However, they code different isoforms as a result of differential RNA processing. For example, in *Drosophila* there appears to be at least three different alternatively spliced transcripts of *dCtBP*[15] (see Chapter by Aihara, Perrone and Nibu). The vertebrate genomes contain two different genes, *CtBP1* and *CtBP2* that code for two highly related proteins. The *CtBP1* gene is located on chromosome 4 of humans and on chromosome 5 of mice. In mammals, the *CtBP1* gene expresses two major transcripts as a result of alternate RNA splicing. These transcripts encode two isoforms of CtBP1, which are identical except for a thirteen amino acid region at the N-terminus (Fig. 1A). The short version of CtBP1 (CtBP1-S) corresponds to an isoform designated as CtBP3/BARS[16] (see Chapter by Spano, Hidalogo Careedo and Corda; the designation CtBP3 has now been changed to CtBP1-S). The transcript for CtBP1-S has an alternate inframe exon (exon 2) in the 5'-region which codes for the N-terminal two amino acids while translation of CtBP1-L is initiated from exon 1 of the shorter transcript (Fig. 1A). A fraction of CtBP1 cDNAs also contains an insertion of a codon for a Ser residue (at position 380 in CtBP1-L and at position 369 in CtBP1-S), which also appears to be the result of alternate RNA processing. The functional significance of the extra Ser residue is not known at present. The CtBP1 proteins are concentrated in the nucleus with significant amounts in the cytosol.