PROTEINS: Structure, Function, and Bioinformatics 63:1069–1083 (2006)

Modeling and Functional Analysis of AEBP1, a Transcriptional Repressor

Peter J. Lyons,[1] Neil R. Mattatall,[2] and Hyo-Sung Ro[1,*]
[1]*Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia, Canada*
[2]*NRC Institute for Marine Biosciences, National Research Council Canada, Halifax, NS, Canada*

ABSTRACT **Adipocyte enhancer binding protein 1 (AEBP1) is a transcriptional repressor of the aP2 gene, which encodes the adipocyte lipid binding protein and is involved in the differentiation of preadipocytes into mature adipocytes. It is an isoform of aortic carboxypeptidase-like protein (ACLP), which is a part of the extracellular matrix. AEBP1 and ACLP contain a conserved carboxypeptidase domain which is critical for the function of AEBP1 as a transcriptional repressor. Homology modeling and multiple alignment of AEBP1 homologues were performed to identify putative domains and critical residues that were then deleted or mutated in mouse AEBP1. Expression of wild-type and mutant AEBP1 proteins in CHO cells was performed, and their function in transcriptional repression was assayed by luciferase assay. All deletion forms of AEBP1 were able to repress transcription driven by the aP2 promoter. The DNA binding domain of AEBP1 was mapped by electrophoretic mobility shift assays to a region of the C-terminus rich in basic residues. However, wild-type AEBP1 was not able to interact strongly with DNA, suggesting that AEBP1 might function predominantly as a corepressor, independent of DNA binding. AEBP1 was also found to interact with Ca^{2+}/calmodulin through this basic region, suggesting another mechanism of functional regulation. Proteins 2006;63:1069–1083.**

Key words: carboxypeptidase; discoidin; homology modeling; DNA binding; calmodulin; transcription

INTRODUCTION

Many genomes have been sequenced, either completely or partially, providing us with a wealth of information about the genes contained within and the proteins encoded by them. Recent research suggests, however, that it is not the genes themselves that are most important in making each organism unique, but the regulation of the activity of each gene.[1,2] There are many mechanisms by which a cell regulates the transcription of a gene. Recent advances in structure determination have given us a clearer picture of the functions and mechanisms of the many factors necessary for transcription initiation and its regulation.[3] In the case of TATA–box-containing promoters, the factor that initiates this sequence of events is the TATA–box binding protein (TBP), which interacts with the minor groove of the DNA at the TATA–box as part of a large multimeric general transcription factor called TFIID. Following the binding of this factor many other general transcription factors (TFIIA, TFIIB, TFIIE, TFIIF, TFIIH, and the Mediator complex) interact with each other and with RNA polymerase II to form a preinitiation complex. Each factor is necessary for a specific function, such as positioning of the complex on the DNA or unwinding of the DNA helix, and each factor can be the target of transcriptional regulators which regulate the initiation of transcription.

In order for gene-specific regulation to occur, each individual gene has specific sequences within its promoter and upstream enhancers that regulate its transcription. Transcription factors bind to these 6 to 12 base-pair–long DNA sequences through sequence-specific DNA binding domains. Most of these DNA-binding domains interact with DNA through an α-helix within a structure such as the helix-turn-helix motif, the homeodomain, the zinc finger motif, the basic leucine zipper motif, or the helix-loop-helix domain.[4] In many cases other parts of the transcription factor protein are also involved in DNA interaction. The protein may also function as a dimer in order to maximize specificity and binding strength.

Following specific DNA binding, different mechanisms are employed by each transcription factor in order to regulate the transcription of a particular gene. A general theme that runs through most transcription factor mechanisms is the modification of the chromatin packaging of DNA in order to recruit, or prevent recruitment, of additional factors and/or basal transcription factors to the site of transcription initiation. Modifications that occur within the chromatin include repositioning whole nucleosomes by ATP-dependent remodeling complexes,[5] replacing individual core histones with variant histones having different functions,[6] and covalent modifications of histones[7–9] by

The Supplementary Materials referred to in this article can be found at http://www.interscience.wiley.com/jpages/0887-3585/suppmat/

Grant sponsor: Natural Sciences and Engineering Research Council of Canada (NSERC); Grant sponsor: Canadian Diabetes Association (CDA); Grant sponsor: Heart and Stroke Foundation of Canada (HSFC); Grant sponsor: Canadian Institute of Health Research (CIHR).

*Correspondence to: Hyo-Sung Ro, Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4H7. E-mail: hsro@dal.ca

Received 12 October 2005; Accepted 16 December 2005

Published online 14 March 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20946

the addition of acetyl,[10] methyl,[11] phosphate,[12] or ubiquitin[13] groups. Transcription is activated or repressed depending on the modification created and the context in which it lies. This context of additional modifications adds to the complexity and flexibility of transcriptional regulation and has been termed the *histone code*.[7,8,14]

Adipocyte enhancer binding protein 1 (AEBP1) was originally characterized as a transcriptional repressor that interacted with a specific upstream promoter element of the aP2 gene termed the *adipocyte enhancer* 1 (AE-1).[15] The aP2 gene expresses adipose fatty acid binding protein (FABP), a protein that is important for the transport of fatty acids within an adipocyte and is upregulated upon adipocyte differentiation.[16] AEBP1 expression is downregulated upon adipocyte differentiation,[15,17] as well as at the terminal stages of osteoblast differentiation.[18] The AEBP1 protein contains three domains, an N-terminal discoidin domain, a central carboxypeptidase (CP) domain, and a C-terminal domain containing regions rich in basic, serine–threonine–proline, and acidic amino acids [Fig. 1(a)].[15] The CP domain was found to be important for the function of AEBP1 as a transcriptional repressor.[15] It was subsequently found that transcriptional repression by AEBP1 is inhibited by interaction with the γ5 subunit of a heterotrimeric G protein within the nuclei of 3T3-L1 preadipocyte cells.[19]

In order to determine the molecular mechanism by which AEBP1 binds DNA and represses transcription, a detailed analysis of the structure of AEBP1 was undertaken. In recent years an increasing number of proteins and protein domains have had their structures solved by X-ray crystallography, giving us a clearer understanding of the structural features that make specific functions possible. Additionally, information provided by recent genomic sequencing efforts has allowed detailed comparisons of protein sequences from many different species in order to determine conserved, and thus critical, residues. Here, we have modeled the structures of the discoidin and CP domains of AEBP1 by homology with recently solved crystal structures. We have performed a bioinformatics analysis of the amino acid sequences of many AEBP1 orthologues obtained by prediction from genomic databases. We have sequentially deleted specific domains of AEBP1 in order to determine the domain necessary for transcriptional repression and DNA binding. Finally, we report another interaction of AEBP1 with Ca^{2+}/calmodulin and characterize the domain responsible for this interaction. While the characterization of the transcriptional repression function of AEBP1 is relevant only to this nuclear protein, the genomic and modeling work presented here will also be of interest to those studying the function of the extracellular AEBP1 isoform, aortic carboxypeptidase-like protein (ACLP),[20–22] as well as the CPX1[23] and CPX2[24] proteins, all of which contain discoidin domains along with inactive carboxypeptidase-like domains [see Fig. 1(a)].

MATERIALS AND METHODS

Homology Modeling and Sequence Analysis

The most critical step in modeling by homology is the alignment of the sequence of the protein to be modeled with that of a suitable template. In this work, suitable template candidates were known from the literature. However, PSI-BLAST[25] was used to search for any additional templates within the PDB library. Using the alignment produced by PSI-BLAST, the program Modeller[26,27] was used to produce a molecular model based on satisfaction of spacial restraints. Model quality was first assessed using ProsaII[28] to detect energy clashes. Procheck,[29] which produces a Ramachandran plot, was also used to analyze the amino acid stereochemistry of the models produced. Finally, SCWRL3[30] was used to predict side-chain placement for those amino acids not identical to the template. When a suitable model was obtained, figures were produced using Molscript[31] and Raster3D.[32]

The majority of protein sequences were obtained from the Ensembl genome database, as well as some from the NCBI databases. Alignments were produced using ClustalW[33] and formatted for figures using GeneDoc.[34] ClustalW also produced a .dnd file, which was used to produce a phylogenetic tree using TreeView.[35]

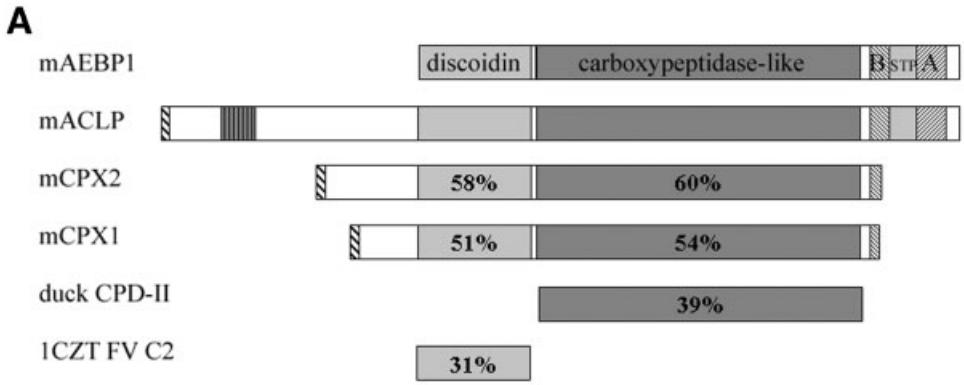
Plasmid DNA Construction and Mutagenesis

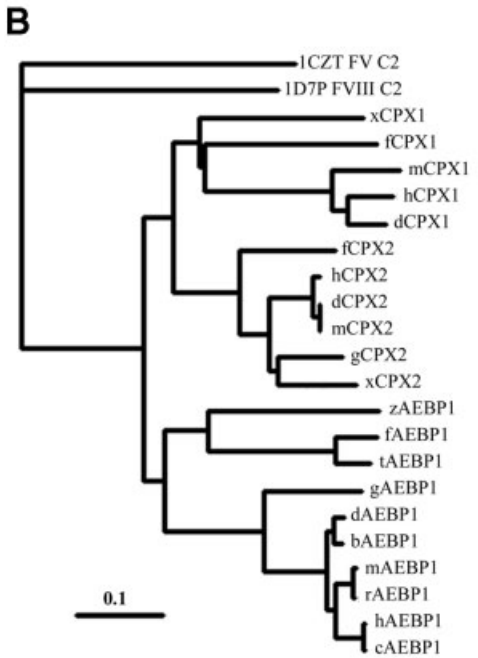
For expression of C-terminally His6-tagged AEBP1 in *E. coli*, DNA encoding the AEBP1 domains of interest was amplified by PCR using PfuTurbo DNA Polymerase (Stratagene) and primers incorporating 5′ NcoI and 3′ HindIII restriction endonuclease sites, sequenced to verify sequence integrity, and inserted into the NcoI and HindIII sites of the pET-21d(+) plasmid (Novagen). Primers AEBP1-1(TATACCATGGAGTCACACCGCATTGA) and AEBP1-2 (TATAAAGCTTGAAGTCCCCAAAGTT CACTG) were used to amplify the insert for pET21d-AEBP1, while primers AEBP1-2 and AEBP1-3 (TATACCATGGTAACTACTGACAGCCTG) were used for pET21d-AEBP1ΔN, AEBP1-1 and AEBP1-4 (TATAAAGCTTGGATCGAGCCAGGATGAA) for pET21d-AEBP1ΔC, AEBP1-3 and AEBP1-4 for pET21d-AEBP1-CP, and AEBP1-1 and AEBP1-5 (TATAAAGCTTATTGAGGCGACGCAGTCG) for pET21d-AEBP1-ΔSTP. The PCR cycle was 30 s at 94°C, followed by 30 cycles of 94°C for 30 s, 58°C for 30 s, and 72°C for 1 min/kb template. Following this cycle, a final extension of 7 min at 72°C was performed.

For expression in mammalian cells, the above pET21d-AEBP1 and derivative plasmids were digested with HindIII and XbaI, blunted with Klenow enzyme, and ligated into pcDNA3.1/myc-HisA similarly digested and blunted. These plasmids expressed the same protein as their parental pET21d plasmid, except lacking a tag and including an unrelated Lys-Leu dipeptide at the C-terminus.

Site-directed mutagenesis was performed using the QuikChange site-directed mutagenesis kit (Stratagene) according to the manufacturer's protocol. The following primers were used:

E239A-sense GGATCCACGGCAATGCGGTGCTAGGCCGAG
E239A-antisense CTCGGCCTAGCACCGCATTGCCGTGGATCC
E363A-sense CATTATTTCCTGGATGGCGAAGAACCCCTTTG


A
mAEBP1
discoidin
carboxypeptidase-like
B
STP
A
mACLP
mCPX2
58%
60%
mCPX1
51%
54%
duck CPD-II
39%
1CZT FV C2
31%


B
1CZT FV C2
1D7P FVIII C2
xCPX1
fCPX1
mCPX1
hCPX1
dCPX1
fCPX2
hCPX2
dCPX2
mCPX2
gCPX2
xCPX2
zAEBP1
fAEBP1
tAEBP1
gAEBP1
dAEBP1
bAEBP1
mAEBP1
rAEBP1
hAEBP1
cAEBP1
0.1

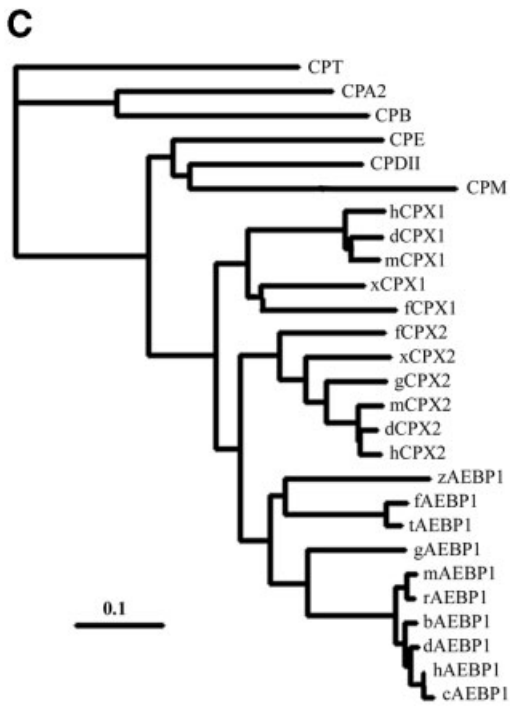

C
CPT
CPA2
CPB
CPE
CPDII
CPM
hCPX1
dCPX1
mCPX1
xCPX1
fCPX1
fCPX2
xCPX2
gCPX2
mCPX2
dCPX2
hCPX2
zAEBP1
fAEBP1
tAEBP1
gAEBP1
mAEBP1
rAEBP1
bAEBP1
dAEBP1
hAEBP1
cAEBP1
0.1

Figure 1.

E363A-antisense CAAAGGGGTTCTTCGCCATCCAGGAAATAATG
N374A-sense GGGTGCAAATCTGGCCGGTGGTGAGCGGCTTG
N374A-antisense CAAGCCGCTCACCACCGGCCAGATTTGCACCC
R388A-sense CCCTATGACATGGCCGCGACACCTAGCCAGG
R388A-antisense CCTGGCTAGGTGTCGCGGCCATGTCATAGGG
R442A-sense GACGGAGCCCTACGCGGGAGGGTGCCAGGC
R442A-antisense GCCTGGCACCCTCCCGCGTAGGGCTCCGTC

The presence of the desired mutation and absence of additional mutations was verified by sequencing. The E239A mutagenesis reaction produced a secondary mutation resulting in the double mutant H236Q/E239A.

AEBP1 Purification

A 500 mL culture of BL21 (DE3) *E. coli* cells, transformed with a pET AEBP1 plasmid, was grown to an OD_{600} of approximately 0.6. Expression of AEBP1 was then induced with 1 to 2 m*M* IPTG for 3.5 h at 250 rpm and at 37°C. Cells were centrifuged at 6000 *g* for 15 min at 4°C and frozen at −20°C at this point unless continuing further.

Cells were resuspended in 25 mL extraction/wash buffer (50 m*M* sodium phosphate, pH 7.0, and 300 m*M* NaCl) and lysed by passing through a French pressure cell twice at 14,000 psi. Immediately prior to lysis, 1 m*M* phenylmethylsulfonyl fluoride (PMSF) was added. Following lysis, extracts were centrifuged at 15,000 *g* for 20 min at 4°C. The supernatant was discarded and the pellet (containing insoluble inclusion bodies) was dissolved in 10 mL denaturing extraction/wash buffer (50 m*M* sodium phosphate, pH 7.0, 300 m*M* NaCl, and 6*M* guanidine-HCl). This was centrifuged again at 15,000 *g* for 20 min at 4°C to clarify.

His6-tagged protein was purified using Talon metal affinity resin (Clontech). A 0.75 mL bed volume of Talon resin was equilibrated twice with 10 mL denaturing extraction/wash buffer by mixing briefly and centrifuging at 700 *g* for 2 min. The clarified protein sample was added to this resin and incubated for 30 min at room temperature on a rocking platform. This was then centrifuged at 700 *g* for 5 min, and the resin washed twice with 10 mL of denaturing extraction/wash buffer, followed by additional washing (5–10 mL) on a 5 mL gravity-flow column until the $OD_{280} < 0.01$. Elution was performed with 3 mL imidizole elution buffer (45 m*M* sodium phosphate pH 7.0, 5.4*M* guanidine-HCl, 270 m*M* NaCl, 300 m*M* imidizole) and 0.5 mL fractions were collected.

Following analysis of protein concentration, fractions with similar concentrations were combined and dialyzed. Dialysis was performed stepwise in 0.5 L volumes of 50 m*M* sodium phosphate, pH 7.5, 100 m*M* NaCl. The buffer during hours 0 to 2 of dialysis contained 4*M* urea, 1 m*M* β-mercaptoethanol, 0.2 m*M* $ZnCl_2$; hours 2 to 4, contained 2*M* urea, 1 m*M* β-mercaptoethanol, 0.1 m*M* $ZnCl_2$; hours 4 to 6 contained 1*M* urea; hours 6 to 8 contained 0.5*M* urea; hours 8 to 10 contained no additives. Dialysis was then continued overnight in a 1.5 L volume of buffer containing 10% glycerol.

Cell Culture and Transfection

Cells were grown in an incubator at 37°C and 5% CO_2. Chinese Hamster Ovary (CHO) cells were grown in DMEM supplemented with 5% FBS, penicillin/streptomycin, and 37 μg/mL L-proline. Transfections were performed in 60 mm dishes or 12-well plates with Polyfect (Qiagen) according to the manufacturer's protocol. Cells were grown an additional 24 to 48 h before harvesting.

Protein Analysis

Cell extracts were prepared in cold radioimmune precipitation buffer (RIPA buffer: 50 m*M* Tris-HCl, pH 7.5, 150 m*M* NaCl, 50 m*M* $Na_4P_2O_7$, 0.25% sodium deoxycholate,

Fig. 1. Phylogenetic comparison of AEBP1, CPX-1, and CPX-2. The Ensembl and NCBI databases were searched for potential orthologues of AEBP1, CPX-1, and CPX-2. These sequences were aligned using ClustalW (see Supplementary Material for a complete alignment). Panel **A** presents a schematic diagram showing similarities in overall makeup of these proteins and percentage identity of individual domains to the homologous domain of ACLP/AEBP1. B, STP, and A indicate basic residue, serine–threonine–proline, and acidic residue rich regions, respectively. Alignments of discoidin domains only (**B**) and of CP domains only (**C**), based on domain boundaries indicated by homology modeling (Figs. 2 and 3), were produced, and phylogenetic trees produced from the .dnd file with Treeview. Branch lengths indicate number of substitutions per site according to the scale at the bottom left. Abbreviations are as follows: m, mouse (*Mus musculus*); r, rat (*Rattus norvegicus*); h, human (*Homo sapiens*); c, chimp (*Pan troglodytes*); d, dog (*Canis familiaris*); b, cow (*Bos taurus*); g, chicken (*Gallus gallus*); f, fugu (*Takifugu rubripes*); t, tetraodon (*Tetraodon nigroviridis*); z, zebrafish (*Danio rerio*); x, *Xenopus tropicalis*. Sequences retrieved from the NCBI databases had the following accession numbers: mAEBP1, CAI25441; hAEBP1, NP_001120; mCPX1, NP_062670; hCPX1, NP_062555; mCPX2, NP_061355; hCPX2, NP_937791. Also found in the NCBI databases was an N-terminally truncated sequence for bAEBP1 (NP_777264), which was made complete with the nucleotide sequence NM_174839, and a C-terminally truncated sequence for cAEBP1 (XM_519550), which was made almost complete by the identification of the C-terminal sequence in the genomic sequence approximately 350 nt 3′ from the last identified exon in Ensembl gene ENSPTRG00000019137. This lack of identification of the last exon was commonly found in many of the following predicted transcripts also. The following Ensembl genes were identified as putative orthologues based on reciprocal BLAST analysis: ENSG00000106624 (hAEBP1), ENSMUSG00000020473 (mAEBP1), ENSRNOG00000013720 (rAEBP1), ENSBTAG00000012237 (bAEBP1), ENSRNOG00000013720 (rAEBP1), ENSCAFG00000002917 (dAEBP1), SINFRUG00000151171 (fAEBP1), ENSDARG00000006901 (zAEBP1), GSTENP00025396001 (tAEBP1), ENSCAFG00000006608 (dCPX1), ENSXETG00000012756 (xCPX1), SINFRUG00000148855 (fCPX1), ENSCAFG00000012718 (dCPX2), ENSGALG00000009693 (gCPX2), ENSXETG00000009571 (xCPX2), SINFRUG00000151238 (fCPX2). The sequence for chicken AEBP1 was assembled from many overlapping EST clones obtained from http://www.chickest.udel.edu (pgf1n.pk006.h24, pgp1n.pk012.m18, pgf2n.pk001.c7, pgm2n.pk007.g13, pgf1n.pk006.o3, pgf1n.pk009.p24) and from http://www.chick.umist.ac.uk (040024.2). The remaining CPs shown in the alignment are human CPA2, cow CPB, CPT from *Thermoactinomyces vulgaris*, the second CP domain of duck CPD, human proCPM (CPM), and human proCPE (CPE). 1CZT FV C2 and 1D7P FVIII C2 indicate the sequence of the corresponding discoidin domain X-ray crystal structure.

0.1% Nonidet P-40, 1 m*M* Na_3VO_4, 1 m*M* NaF) containing 1 m*M* PMSF, 5 m*M* EDTA, 5 m*M* EGTA, and Complete Protease Inhibitor Cocktail (Roche).

Proteins were resolved on 8.5% SDS polyacrylamide gels according to established methods. Following SDS-PAGE, proteins were transferred to membrane by Western blotting or gels were washed three times with water and stained with GelCode Blue (Pierce) for 1 h to overnight. Destaining was performed with water until complete.

In order to analyze resolved proteins by immunoblotting, proteins were transferred to supported nitrocellulose membrane (Bio-Rad) using a Trans-Blot SD Semi-dry Transfer Cell (Bio-Rad) for 50 min at approximately 400 mA in transfer buffer (48 m*M* Tris-HCl, 39 m*M* glycine, 3.75 mL 10% SDS/L, and 20% methanol). The membrane was blocked with 5% nonfat milk powder in TBST for 1 h at room temperature, followed by incubation with primary antibody, diluted in 5% milk in TBST at 4°C overnight.

Transcription Assays

Cells were transfected in a 12-well dish. Forty-eight hours after transfection, cells were washed twice with PBS and lysed in 150 μL Passive Lysis Buffer (Promega). Fifty microliters of lysate was transferred to a 96-well plate. Luciferase activity was measured by injection of 50 μL Luciferase Reagent (Promega) and measurement of luminescence by a FLUOstar Galaxy luminometer (BMG Labtechnologies).

β-Galactosidase activity was assayed in order to normalize for transfection efficiency. Thirty microliters of extract was incubated with 3 μL Mg^{2+} buffer (10 μL 1*M* $MgCl_2$, 35 μL β-mercaptoethanol, 55 μL water), 66 μL ONPG (4 mg/mL in 0.1*M* $NaHPO_4/Na_2PO_4$, pH 7.4), and 201 μL 0.1*M* $NaHPO_4/Na_2PO_4$ buffer (pH 7.4) until color change was sufficient. Reactions were stopped by the addition of 0.5 mL 1*M* Na_2CO_3, and enzyme activity measured spectrophotometrically at A_{410}.

Electrophoretic Mobility Shift Assay (EMSA)

AE-1 duplex oligonucleotide was radiolabeled with [α-^{32}P]dATP by Klenow fill-in. Recombinant proteins were incubated with DNA probe for 25 min at room temperature in binding buffer (10 m*M* Tris, pH 7.5, 10 m*M* KCl, 5 m*M* MgCl2, 1 m*M* DTT, 2.5 % glycerol). Samples were resolved on 5% 0.25× TBE mini-gels (2.5 mL 29:1 acrylamide: bisacrylamide, 375 μL 10× TBE, 95 μL 10% APS, 10 μL TEMED, 12.0 mL water) which were then dried and exposed to film.

Calmodulin-Agarose Pull-Down Assays

Thirty microliters calmodulin-agarose (Sigma) was washed with 0.5 mL HBS. One hundread fifty microliters HBS/0.1% NP40 was added along with 5 μg recombinant AEBP1 and 2 m*M* $CaCl_2$ or 10 m*M* EGTA and incubated for 45 min at room temperature on a rocking platform. The beads were then washed four times with 1 mL HBS/0.1% NP40, 20 μL SDS-PAGE sample buffer was added, and bound proteins were resolved on an 8.5% SDS-PAGE gel. Proteins were visualized by staining with GelCode Blue (Pierce).

RESULTS AND DISCUSSION

Phylogenetic Analysis of AEBP1, CPX-1, and CPX-2

Genomic databases were searched for putative orthologues of mouse AEBP1. These searches identified a number of AEBP1 candidates from vertebrate species. No potential AEBP1 orthologues were identified in nonvertebrate organisms such as *C. elegans*, *S. cerevisiae*, or *D. melanogaster*. Orthologues of two close relatives of AEBP1, CPX-1 and CPX-2, were also identified, all of which contain conserved discoidin domains as well as carboxypeptidase domains [Fig. 1(a)]. Due in part to sequence divergence, it was not possible to conclusively identify the exons encoding the amino acids N-terminal to the discoidin domain in many orthologues of ACLP, CPX-1, and CPX-2. Because ACLP, CPX-1, and CPX-2 have very similar discoidin and CP domains, phylogenetic analysis was performed to confirm the identity of each protein. This analysis indicated that orthologues of each protein fell into one of three distinct groups defined as ACLP, CPX-1, and CPX-2 as predicted, even when only the discoidin or CP domain sequences were used in the analysis [Fig. 1(b,c)]. In fact, ACLP, CPX-1, and CPX-2 form their own subgroup when compared with other members of the N/E carboxypeptidase subfamily [Fig. 1(c)]. In addition to differences in local sequence, large-scale features distinguish AEBP1/ACLP from CPX-1 and CPX-2. Both CPX-1 and CPX-2 have large N-terminal extensions, distinct from that present in an alternatively spliced isoform of AEBP1, ACLP, which also has an N-terminal extension. CPX-1 and -2 lack the large C-terminal domain of AEBP1, but instead have a short 10-to-15 residue basic stretch at their C-termini. Finally, AEBP1 contains two large internal loops within the CP domain not present in most other carboxypeptidases, the second of which is lacking in CPX-1 and CPX-2. However, this second loop is also not conserved in nonmammalian AEBP1/ACLP (see Supplementary Material).

Structural Modeling of the Discoidin Domain of AEBP1

The structures of two proteins with homology to the discoidin domains of AEBP1 and CPX-1 and -2 have recently been solved by X-ray crystallography.[36,37] These domains, the C2 domains of human coagulation factors V (PDB ID 1CZT) and VIII (PDB ID 1D7P), are composed of a β-barrel framework from which three spikes protrude. These spikes are involved in membrane interaction through hydrophobic residues at the apices of two spikes which embed themselves in the hydrophobic membrane core.[36,37] Membrane phosphatidylserine head groups also make interactions with the grooves between the spikes of the C2 domains, while the negatively charged membrane surface interacts with many basic residues forming a ring around the C2 domain. The C2 domains of both coagulation factors V and VIII share considerable sequence similarity (~ 50%) with the discoidin domain of AEBP1.

In order to determine the three-dimensional structure of the discoidin domain of AEBP1, we employed homology modeling using the program Modeller.[27] The modeled structure of the discoidin domain of AEBP1 reveals several interesting points (Fig. 2). First, two cysteines, C45 and C143, are found in very close three-dimensional proximity to each other. These cysteines are not found in the template molecules 1CZT or 1D7P but are conserved in all orthologues of AEBP1, CPX-1 and CPX-2 found to date (see Supplementary Material). The proximity of these cysteines to each other in the tertiary structure, as well as total conservation, suggests that they may form a disulfide bond contributing to the stability of the domain, and so we have constrained the model to include this disulfide bond [Fig. 2(b)]. Another conserved cysteine is present in this model at amino acid 151. A homologous cysteine is present in the discoidin domains of both Factors V and VIII and is disulfide bonded to a cysteine located at the extreme N-terminus of the domain [see Fig. 2(a)]. This cysteine is lacking in AEBP1, which is missing the seven N-terminal amino acids of the discoidin domain. However, the sequence for ACLP (see Supplementary Material) indicates a conserved cysteine in a similar location, shifted by one amino acid, and is likely disulfide bonded with the cysteine at the C-terminus of the discoidin domain. These predicted disulfide bonds may serve to hold the N-terminus of the discoidin domain packed tightly against the rest of the discoidin domain. The N-terminal 45 amino acids of this domain contain a high proportion of charged residues, with seven basic amino acids out of a total of 16 in the domain, while the opposite face of the DLD domain is largely hydrophobic and non-charged [see Fig. 2(c)].

The model of the DLD domain of AEBP1 indicates several spikes with similarity to the spikes of the template molecules involved in membrane interactions [Fig. 2(b)]. These residues in the Factor V C2 domain (1CZT), used as a template for AEBP1, are W27 and W28 (spike 1), R44 (spike 2), and L80 [spike 3; see Fig. 2(a)]. The Factor VIII C2 domain also has three spikes, capped by M29 and F30 (spike 1), R45 (spike 2), and L81 and L82 [spike 3; see Fig. 2(a)]. These domains lie such that the hydrophobic spikes insert into the membrane core, while the basic spikes and surrounding basic residues interact with the negatively charged surface of the membrane. AEBP1 does not contain hydrophobic residues at the tip of the spike 1 equivalent, but instead contains an arginine and a histidine similar to spike 2 of the coagulation factor C2 domains, the histidine being strictly conserved (see Supplementary Material). A potential surface exposed loop is suggested by the sequence ^{33}GANEDDYYDG^{42} which contains a well-conserved seven-amino-acid loop not present in the template structures. This loop contains four acidic amino acids and two tyrosines, the tyrosines being modeled at the apex of the loop. In the model this loop is placed so that the tyrosines at the apex are inserted into the central cavity of the domain. However, the sequence similarity with the template molecule in this region of the domain is relatively weak. If this loop were to face the opposite direction, these tyrosines would be on the surface of the protein, possibly functionally replacing the hydrophobic residues found at the tip of spike 1 of the C2 domain of coagulation factor V.[36] AEBP1 Ser76 and Ile77 are structurally similar to those residues found at the tips of spike 3 in the coagulation factor C2 domains. Ile77 is conserved in most AEBP1 orthologues identified except Fugu and Tetraodon (see Supplementary Material). It is also interesting to note that CPX-1 and -2 residues equivalent to these spike-capping residues are consistently large hydrophobic residues such as tyrosine, phenylalaline, and tryptophan (Supplementary Material). Altogether it appears that the discoidin domain of AEBP1/ACLP, as well as that of CPX-1 and -2, is structurally similar to coagulation factor V and VIII C2 domains and may function in a similar manner in the interaction with membrane surfaces. CPX-1 and -2 are suggested to be extracellular proteins, and may interact with the outer membrane surfaces of cells. Immunoreactivity towards an α-AEBP1 antibody has also been detected at the cell membrane of hyperplastic mammary epithelial cells using immunohistochemical techniques (unpublished data).

Structural Modeling of the Carboxypeptidase Domain of AEBP1

The central half of the AEBP1 protein consists of a carboxypeptidase (CP)-like domain. This domain is a member of the regulatory carboxypeptidase family due to its sequence and structural similarities. Recently, the first X-ray crystal structure of a member of the regulatory CP family, CPD-II, has been solved.[38] Carboxypeptidase D (CPD) contains three tandem repeats of a CP domain. The first domain is enzymatically more efficient towards C-terminal arginine, the second domain is more efficient towards C-terminal lysine, and the third domain is inactive.[39] Gomis-Ruth and colleagues[38] determined the three-dimensional structure of the second domain of CPD, the more highly conserved of the two active domains, as well as the CPD-II domain in complex with an inhibitor, GEMSA.[40] Duck CPD-II has 40% sequence identity and 54% similarity to the CP domain of AEBP1.

The construction of CP domain models (Fig. 3), both with and without a peptidomimetic inhibitor, allows us to make some predictions regarding the structure and function of AEBP1. The CP domain of AEBP1 is predicted to contain three disulfide bonds based on proximity of cysteine residues to each other, only one of which is present in the CPD structure. These disulfide bonds are Cys192–Cys254, Cys445–Cys488, and Cys578–Cys589. The first disulfide bond may link together two helices on the surface of the domain, and may serve to restrain the placement of the N-terminal discoidin domain with respect to the CP domain. This disulfide bond is also present in the homologous CPE sequence, based on biochemical measurements.[40] The second predicted disulfide bond, which is also present in the CPD-II structure, links a loop on the surface of the domain with the core of the domain, thus restraining the motion of this loop. The third predicted disulfide bond is located in the C-terminal subdomain. The structure of this subdomain in CPD-II is maintained by a buried cluster of

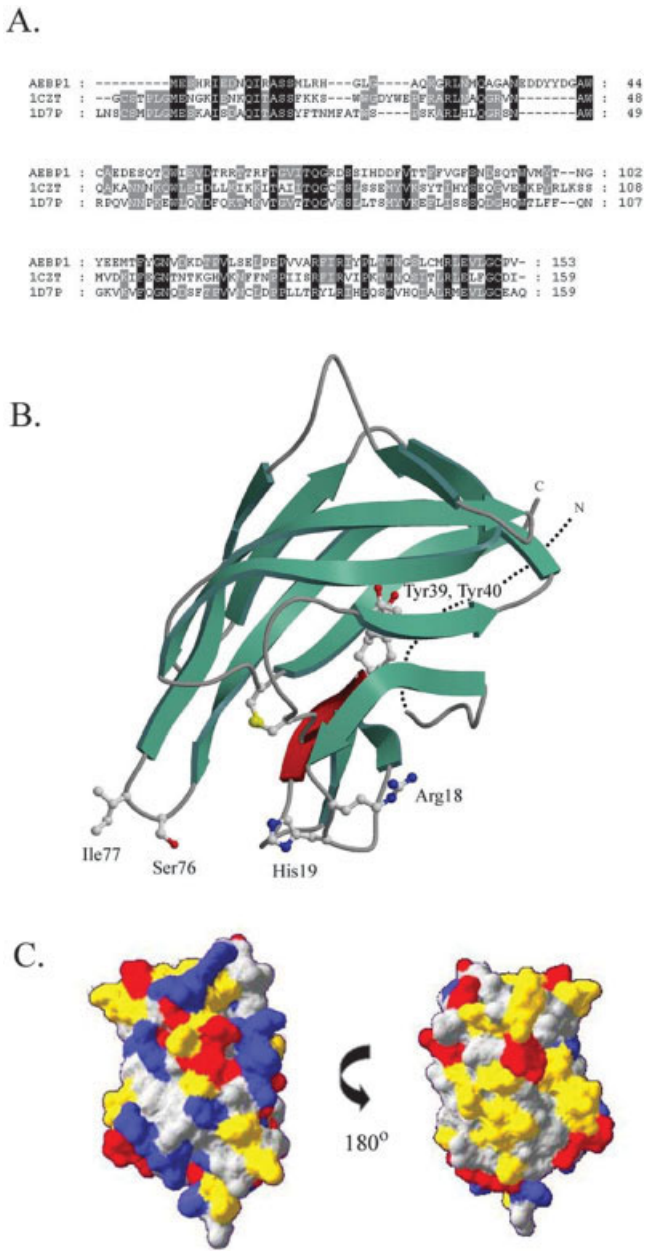


Fig. 2. The DLD domain of AEBP1 is a β-barrel with two distinct faces. (**A**) The amino acid sequence of the discoidin-like domain (DLD) of AEBP1 was aligned with the amino acid sequences of two homologous domains for which X-ray crystal structures have been solved — 1CZT (coagulation factor V C2 domain) and 1D7P (coagulation factor VIII C2 domain). The alignment was performed with ClustalW followed by shading with GeneDoc. A black background indicates identical residues and a grey background indicates chemically similar residues. (**B**) The homology model of the discoidin domain of AEBP1 is shown. One predicted disulfide bond is indicated by a ball and stick structure, connecting a highly charged N-terminal subdomain to the remainder of the domain. β-strands are indicated in green and α-helices in red. Amino acids 36 to 42, which could potentially be solvent exposed, forming a spike, are shown in red. All residues potentially forming the tips of spikes are shown as ball-and-stick structures and labeled. The N- and C-termini are labeled as such, with the seven N-terminal amino acids of the discoidin domain that are not present in AEBP1 shown as a dotted line. The N-terminal subdomain contains a large percentage of basic amino acids, thus resulting in a polarization of the DLD domain of AEBP1 with one face being highly positively charged (**C**) and the other face being uncharged/hydrophobic (**C**). Basic residues are indicated in blue, acidic residues in red, polar residues in yellow, and nonpolar residues in white. Panel (**B**) was created with Molscript/Raster3D and (**C**) with Swiss PDB Viewer.

hydrophobic amino acids running the length of the subdomain.[38] Interestingly, the final two β-strands of this subdomain lack the majority of these hydrophobic residues in AEBP1. Instead, this region of the subdomain seems to be held together through this third putative disulfide bond functioning in lieu of the hydrophobic cluster. It is interesting to note the rarity of disulfide bond formation within the cytosol and nucleus. It is generally understood that disulfide bonds form within the endoplasmic reticulum in proteins destined for the secretory pathway, as are ACLP, CPX1, and CPX2. However, some proteins form disulfide bonds within the cytosol as part of their catalytic cycle while others contain cysteine residues that are susceptible to oxidation under both normal and oxidative stress conditions.[41] AEBP1 may fall into this rare class or may lack disulfide bonds otherwise found in ACLP.

The putative active site of AEBP1 has several distinctive features. Two loops which were deleted from the AEBP1 sequence in order to create the AEBP1 CP model most likely form part of the funnel leading to the putative active site, as indicated in Figure 3(b). This is also the case in CPE[40] and CPB, which also have large loops in the same place as the first loop of AEBP1, although not as large. The active site itself is lacking many residues critical for CP enzymatic activity in most carboxypeptidases. While two zinc ligands, His236 and Glu239, are conserved in AEBP1, the remaining histidine zinc ligand and glutamate general base are not, but are modified to Asn374 and Tyr485, respectively. While these are not normally metal coordinating or catalytic residues in carboxypeptidases, there is some precedent for amino acids with carboxamide sidechains (asparagine and glutamine) and tyrosine being involved in these functions. The crystal structure for phosphomannose isomerase has indicated that the catalytic zinc ion is coordinated in a distorted trigonal bipyramidal arrangement by two histidines, a glutamate, a water molecule, and the amide oxygen of a glutamine.[42] The crystal structure of isopenicillin N synthase (IPNS) also indicated a glutamine as a metal coordinating ligand, although it was later found that this residue was not necessary for IPNS catalytic activity.[43,44] An experiment investigating the coordination of the catalytic zinc ion within carbonic anhydrase II (CAII) found that substitution of zinc-coordinating histidines with asparagine or glutamine retained metal coordination ability, although with diminished affinity.[45] Tyrosine is involved in zinc coordination in the astacin peptidase, a member of the metzincin family of metalloproteases.[46] In addition to its function in zinc coordination, this tyrosine also plays a role in transition state stabilization similar to Tyr248 of CPA. Another protein of the gluzincin family of peptidases, leukotriene A4 hydrolase (LTA4H), contains a tyrosine at position 383 which, although not necessary for zinc coordination, is critically important as a proton donor in the catalytic mechanism of peptide cleavage.[47,48] Figure 3(c) illustrates the potential coordination of this zinc ion in AEBP1, with both asparagine 374 and tyrosine 485 (through a water molecule) playing a role in zinc coordination.

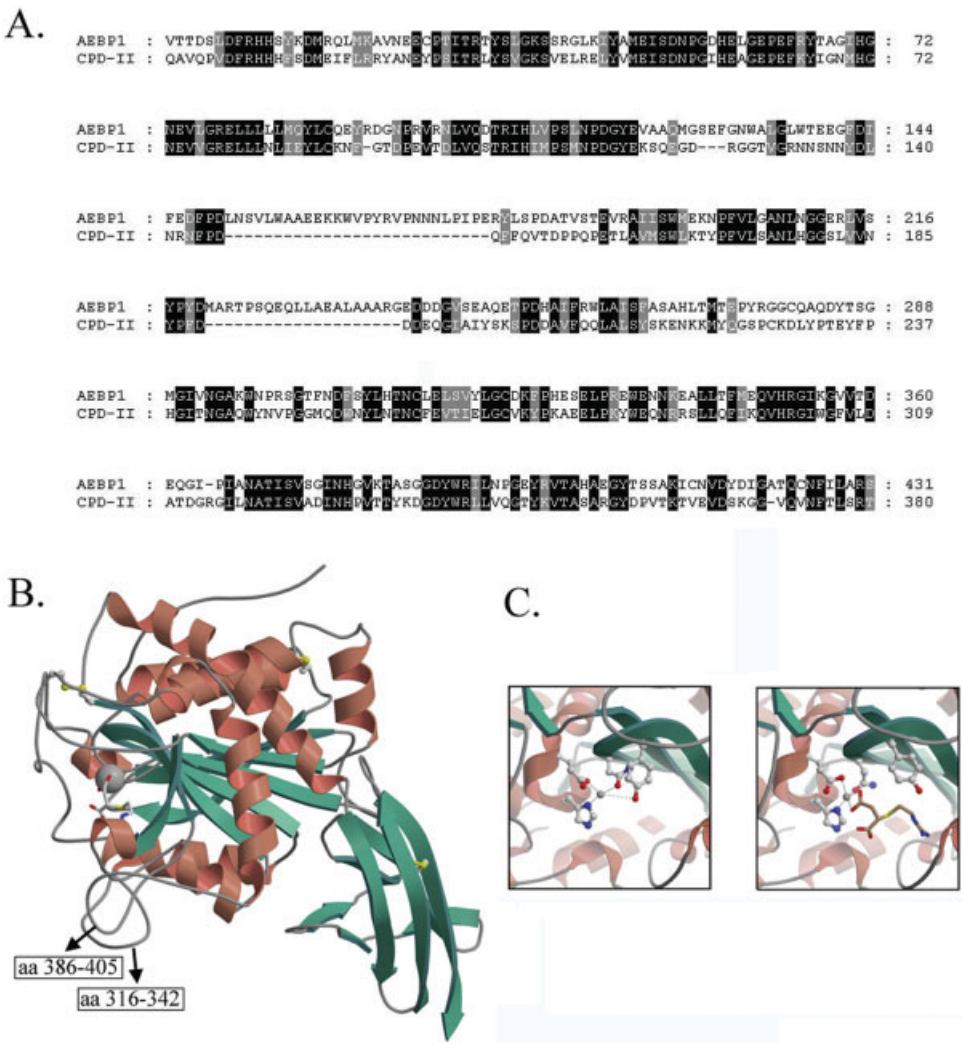


Fig. 3. The CP domain of AEBP1 as modeled by homology against duck carboxypeptidase D. (**A**) Mouse AEBP1 was aligned using ClustalW with the protein sequence for the duck CPD-II domain as used for the determined X-ray crystal structure, 1QMU. A black background indicates identity while a grey background indicates similarity. Shading was produced using GeneDoc. (**B**) The CP domain of AEBP1 has two subdomains, a larger subdomain composed of a central eight-stranded β-sheet surrounded by nine α-helices. Three predicted disulfide bonds are indicated by ball and stick structures, the modeled inhibitor, GEMSA, is indicated by a stick structure, and the putative catalytic zinc ion is indicated by a grey sphere. β-strands are shown in green and α-helices in red. The putative positions of two loops that were not modeled are indicated. (**C**) The putative active site of AEBP1 is shown, with lines indicating bonds coordinating the catalytic zinc. The left panel shows the putative active site of AEBP1 as modeled against 1QMU. The right panel shows the putative active site, as modeled against 1H8L and thus including the inhibitor, GEMSA. This figure was created with Molscript/Raster3D. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

The possibility remains that N374 and Y485 of AEBP1 may not be involved in zinc coordination or the catalytic mechanism. In fact, the multiple alignment of AEBP1 orthologues (see Supplementary Material) indicates that asparagine 374 is substituted with a glutamine in all nonmammalian AEBP1 orthologues and tyrosine 485 is replaced with a phenylalanine in cow and fish orthologues, suggesting that these residues are not critical. While asparagine and glutamine both contain the same functional group, phenylalanine lacks the hydroxyl group proposed to be involved in zinc coordination. In the protease astacin, which contains a tyrosine at position 149 involved in zinc coordination, mutation of this residue to phenylalanine abolished all but 2.5% of wild-type enzyme activity.[46] There has been some controversy regarding whether or not AEBP1 has CP activity. Early work from our laboratory detected low CP activity towards traditional CPB substrates.[15,49] However, other labs have not been able to detect this low activity,[21,50,51] and recently we also have had difficulty detecting CP activity using standard substrates (unpublished data).

Other CP residues have been identified as having important roles in carboxypeptidase substrate binding. CPD and all members of the N/E family contain an aspartate

(position 192 in CPD-II and position 385 in AEBP1) which binds the side chain of C-terminal basic amino acids. This is conserved in all AEBP1 orthologues (see Supplementary Material). The crystal structure of CPD-II bound to GEMSA[40] indicates the inhibitor carboxylate group of GEMSA, which mimics a peptide substrate C-terminus bound by R135, N144, and R145. These residues are substituted in mouse AEBP1 by L301, F310, and E311, respectively. These residues are conserved in mammalian AEBP1 (although F310 is sometimes a tyrosine), but not in nonmammalian orthologues identified thus far. A peptidase from *B. sphaericus* also lacks two of these specific residues, R135 being substituted by an asparigine and R145 being substituted by aspartate.[52–54] This peptidase is able to hydrolyse C-terminal *meso*-diaminopimelic acid, which contains a free amino group in addition to the carboxyl group. This suggests that AEBP1 may have specificity towards amino-terminal substrate residues.

Expression of AEBP1 Deletion and Mutation Versions in CHO cells

In order to characterize the roles of the individual domains of AEBP1 in the structure and function of AEBP1, plasmids expressing various deletion forms of AEBP1 were constructed. These plasmids were transfected into CHO cells and resulted in overexpression of full-length and all truncated versions of AEBP1 [Fig. 4(b)]. A polyclonal anti-AEBP1 antibody was used to detect the overexpressed protein. Because the mutant versions of AEBP1 may lack some of the epitopes recognized by this antibody, we cannot be confident that the band intensities truly correlate with protein abundance. However, in the event that the deletions and mutants are recognized by the anti-AEBP1 antibody to a similar extent as wild-type AEBP1, we might make several conclusions. The AEBP1ΔN mutant protein consistently exhibited lower protein levels than the other versions of AEBP1. Wild-type, ΔC, and ΔSTP AEBP1 protein levels were typically similar, while AEBP1 CP levels were often higher than the levels of the other proteins detected by the anti-AEBP1 antibody. When full-length AEBP1 was overexpressed from pcDNA–AEBP1, a single major species with an apparent molecular weight of 85 kDa was observed along with a minor band of about 60 kDa. It is thought that this minor band is due to an N-terminal ~ 25 kDa truncation, as a similar band was observed upon overexpression of a C-terminally myc/His-tagged version of AEBP1, although shifted upwards by an amount consistent with the size of the C-terminal tag [Fig. 4(c), lanes 10 and 11].

Several site-directed mutants and smaller deletions of AEBP1 were also made based on the homology modeling and phylogenetic analysis described in the previous section [Fig. 4(c)]. Two large loops within the CP domain of AEBP1 that are not present in many related CPs were deleted to produce AEBP1 Δ316–342 and Δ386–405 deletions. Three residues, H236, E239, and N374, equivalent to known zinc coordinating residues in related CPs, were mutated to glutamine or alanine in a double mutant, H236Q/E239A, and a single mutant, N374A. These proteins were not expected to exhibit any CP activity due to a lack of zinc coordination. E363 was predicted to form a salt bridge with H476, both residues being conserved in all identified sequences. This predicted salt bridge would be disrupted upon mutation of E363 to an alanine, possibly destabilizing the protein structure. R388 is the only absolutely conserved residue in the residue 386 to 405 loops and predicted through analysis of the homology model to possibly function in substrate binding. R442 was also thought to be a possible substrate carboxyl-terminal binding candidate residue, in the event that E311 does not function in N-terminal group interaction but is substituted by an alternative C-terminal binding group. These were expressed in CHO cells and analyzed by immunoblotting. All but two of the deleted or mutant proteins exhibited levels similar to wild-type AEBP1. Δ386–405 AEBP1 and R388A both had much lower steady state levels than WT and other mutant forms of AEBP1. R388 is the only residue in the amino acids 386 to 405 loop that is conserved in all AEBP1 orthologues analyzed to date, and therefore the lack of this one residue may be responsible for the decreased levels of Δ386–405 AEBP1 also. It may be that R388 is important for maintaining the tertiary structure of AEBP1 or for some other aspect of protein stability.

Transcriptional Repression Ability of AEBP1 Truncation Forms

To determine the ability of these deletion and mutant forms of AEBP1 to repress transcription, they were expressed in CHO cells and analyzed for their ability to repress transcription of a luciferase reporter gene controlled by the aP2 promoter (Fig. 5). A β-galactosidase–expressing plasmid driven by the CMV promoter was also cotransfected as an internal control for transfection efficiency and was expected to be unaffected by expression of AEBP1. When the luciferase assay data was normalized for transfection efficiency with β-galactosidase assay data, both ΔN and ΔC AEBP1 repressed transcription at least as well as, if not better than, WT AEBP1 [Fig. 5(a)]. However, it was unexpectedly found that β-galactosidase activity also decreased as cells were transfected with higher levels of plasmid expressing AEBP1 WT or AEBP1 ΔC [Fig. 5(c)]. This decrease was not as great as the reduction in aP2-regulated luciferase gene expression [Fig. 5(b)], resulting in a net repression of the aP2 promoter upon normalization, although attenuated compared to the unnormalized luciferase data. ΔN AEBP1 had little effect on β-galactosidase activity, although the same amount of DNA was used under the same conditions. ΔN AEBP1 also had a somewhat reduced effect on aP2-driven luciferase expression when compared to WT and ΔC AEBP1. The dose-dependent decrease in β-galactosidase activity only when increasing amounts of WT AEBP1 or ΔC AEBP1 were expressed suggests that this reduction is due to repression by AEBP1 and not merely due to a decrease in transfection efficiency. The absence of any effect on β-galactosidase expression upon deletion of the amino terminal portion of AEBP1 suggests that this deletion may result in impaired recognition of the CMV promoter that drives expression of the

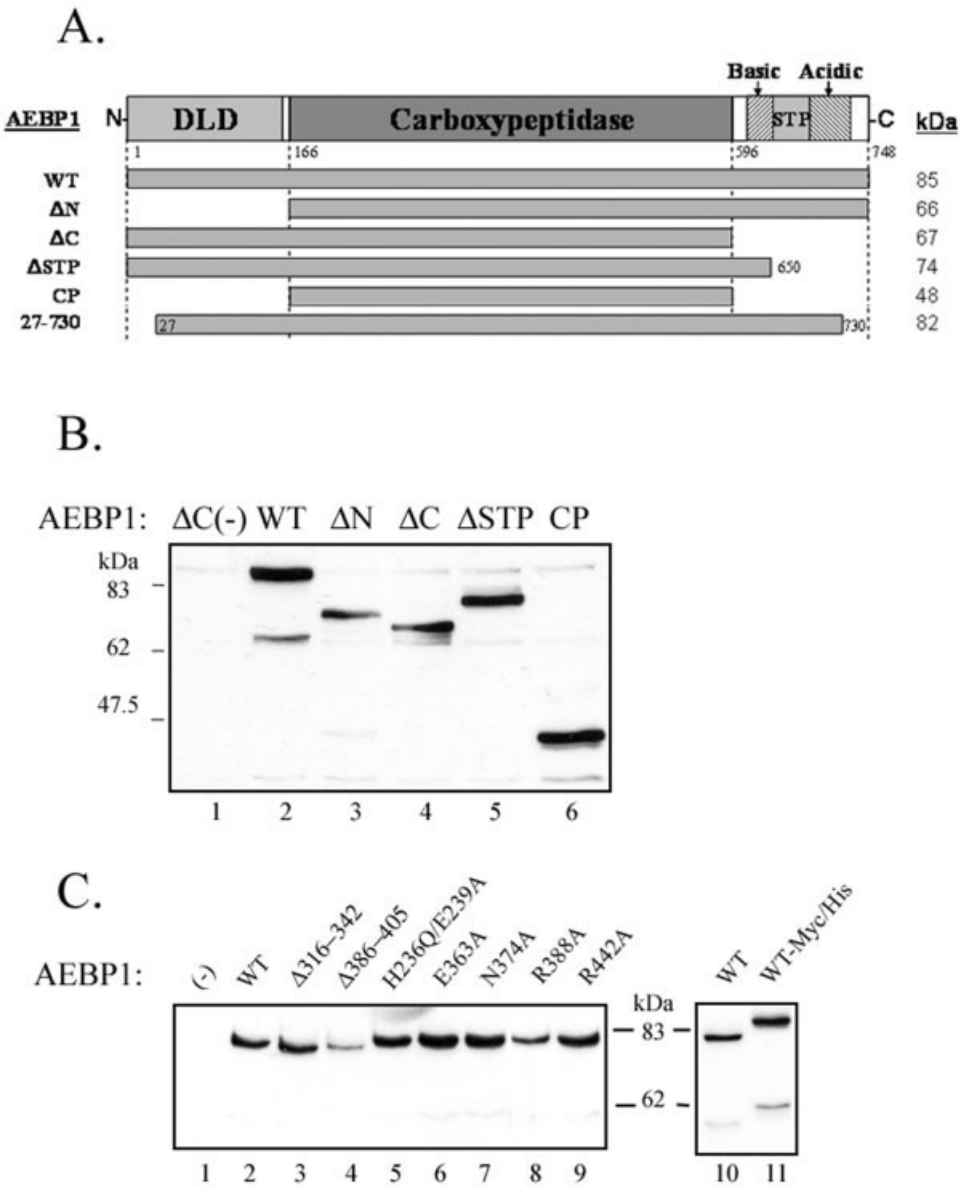


Fig. 4. Expression of wild-type and mutant versions of AEBP1 in CHO cells. (**A**) A diagram is shown illustrating the domain structure of AEBP1 and the various truncation versions of AEBP1 used in this study along with their calculated molecular weights. DLD, discoidin-like domain; basic, basic amino acid-rich region; STP, serine–threonine–proline amino acid rich region; acidic, acidic amino acid rich region. AEBP1 and derivative deletions (**B**) and mutants (**C**) were expressed from a pcDNA based vector in CHO cells. AEBP1 shown in lane 11 of (**C**) was C-terminally fused with a Myc/His tag, while all other proteins were untagged. As negative controls, cells were transfected with a pcDNA3.1 plasmid in which the AEBP1ΔC ORF was inserted in the opposite orientation [ΔC(−)] or empty vector (−). Total protein was extracted from the transfected cells and equal amounts were separated by SDS-PAGE and immunoblotted with an anti-AEBP1 polyclonal antibody.

LacZ gene on the expression plasmid, while still allowing regulation of the aP2 promoter. Alternatively, the lower levels of ΔN AEBP1 might result in weaker transcriptional repression, with varying effects dependent on the assay used.

In contrast to the effects of ΔN AEBP1, WT and ΔC AEBP1 were able to reduce expression from both promoters. Similar results were obtained when empty expression vector was transfected as a control and upon transfection of plasmid from different plasmid preparations (results not shown), suggesting that the results are not due to a DNA effect or due to plasmid impurities affecting transfection efficiency. The ability of ΔC AEBP1 to repress transcription from both the aP2 and CMV promoters was essentially the same as the ability of WT AEBP1 to repress transcription from these promoters, even when expressing very low levels of protein. This suggests that AEBP1 may not interact with DNA through its C-terminus or that DNA binding through the C-terminus of AEBP1 may not play an important role in transcriptional repression in CHO cells.

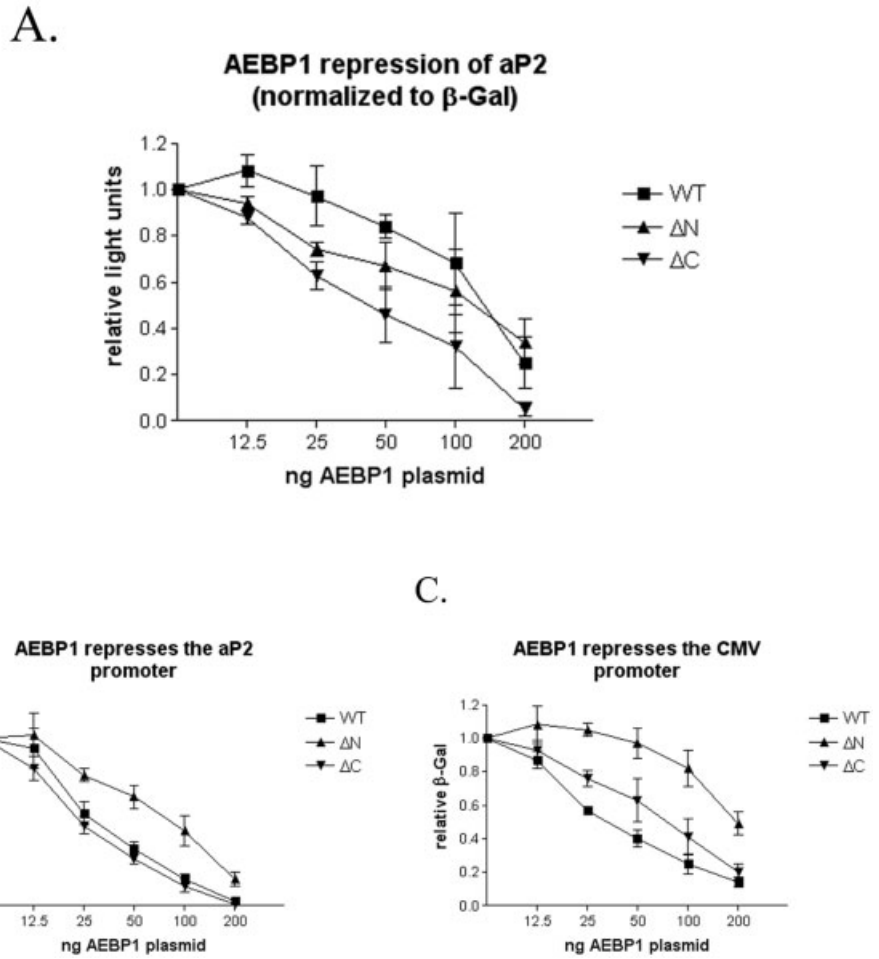


Fig. 5. N- and C-terminally truncated AEBP1 repress transcription from the aP2 promoter. CHO cells were transfected with three plasmids, one containing the aP2-luciferase reporter gene, the second expressing β-galactosidase (pHermes-LacZ), and the third expressing AEBP1 WT or the ΔN or ΔC versions of AEBP1. The amount of the plasmid expressing AEBP1 or truncations was varied as indicated and varying amounts of pcDNA empty vector was added to ensure that the total amount of DNA was equal in all transfections. Cells were lysed and transcription of the reporter plasmid measured as light produced normalized to β-galactosidase activity (**A**). The unnormalized results from the luciferase assay (**B**) as well as the β-galactosidase assay (**C**) are also shown. Results are averages (± SD) from four transfections.

DNA Binding by AEBP1

In order to better characterize the ability of AEBP1 to interact with AE-1 DNA oligonucleotide and the relationship of this DNA binding to transcriptional repression by AEBP1, full-length and domain-specific deletions of AEBP1 were used in electrophoretic mobility shift assays (Fig. 6). A signal representing the radiolabeled AE-1 DNA duplex bound with AEBP1 protein or a deletion form of AEBP1 was observed for all proteins. WT AEBP1 consistently exhibited a very weak signal representing protein bound to a small fraction of total radiolabeled DNA. ΔN and ΔC (data not shown) AEBP1 also bound DNA weakly, although stronger than WT AEBP1. AEBP1(27–730) and ΔSTP AEBP1 were both much more efficient at DNA binding, with ΔSTP by far having the greatest ability to bind DNA. A second complex with lower mobility was also observed upon interaction of ΔSTP with AE-1 DNA, suggestive of protein dimerization. The weak binding of the ΔC deletion (results not shown) yet very strong binding of the ΔSTP deletion suggests that the primary DNA binding domain of AEBP1 is the basic region within the C-terminus. It appears that both the ΔSTP and the AEBP1(27–730) forms of AEBP1 lack regions that may normally regulate the ability of the protein to interact with DNA.

Table I presents a summary of all in vitro DNA binding data for AEBP1. Only two forms of AEBP1, AEBP1(27–730) and AEBP1ΔSTP, are able to interact strongly with DNA, and both lack the extreme C-terminus but still contain the basic helix within the C-terminal domain. In contrast, all forms of AEBP1 that contain the C-terminal hydrophobic region are unable to bind DNA. This data suggests that the presence of the C-terminal hydrophobic stretch of the C-terminal domain prevents interaction with DNA that could otherwise occur through the basic helix within the C-terminus. The C-terminal 22 amino acids of

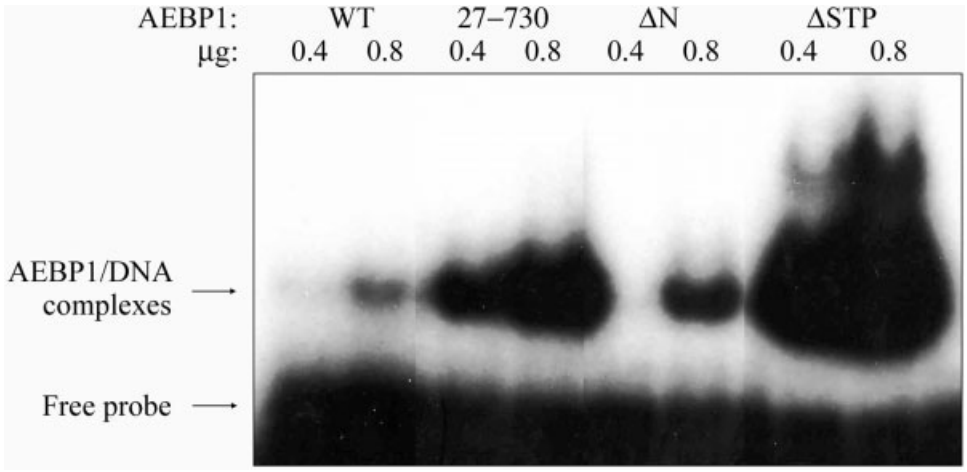


Fig. 6. Wild-type AEBP1 binds DNA weakly through its basic region. Electrophoretic mobility shift assays (EMSA) were performed with radiolabeled AE-1 duplex oligonucleotide and the indicated amounts of *E. coli* expressed and purified AEBP1 or its deletion derivatives.

TABLE 1. Ability of AEBP1 and Deletions to Interact with AE-1 Duplex DNA

AEBP1 Protein Used in In Vitro EMSA	Source of Information	Strength of DNA Binding
WT	This work (Fig. 6)	−
(31–748)	Ref.[20]	−
(123–748)	Ref.[57]	−
(27–730)	This work (Fig. 6)	+++
ΔSty (27–543)	(Muise and Ro, unpublished results)	−
ΔN	This work (Fig. 6)	+
ΔC	Data not shown	+
ΔSTP	This work (Fig. 6)	++++
CP	Data not shown	+

AEBP1 are largely uncharged except for three negatively charged residues and stand in striking contrast to the long stretch of negatively charged glutamate located nearby. In fact, a consensus sequence within this hydrophobic stretch [733-(Y/F)PΦxxxxxxYT(V/I)(D/N)-744, in which the numbering is taken from the mouse sequence and Φ indicates a hydrophobic amino acid] can be identified with several invariant residues in all orthologues (see Supplementary Material). This is unlike the majority of the C-terminus of AEBP1, which is not well conserved. The hydrophobic nature of the C-terminus suggests that it may be inserted into the hydrophobic interior of the protein. Attempts to purify soluble C-terminally His6-tagged AEBP1 using a metal affinity column support the idea that the C-terminus is not surface exposed, as very little (~ 2%–5% of the total) wild-type AEBP1 was able to bind the column through the C-terminal His6 tag (results not shown). On the other hand, it was found that approximately 50% of soluble AEBP1–ΔC and AEBP1–ΔSTP, both lacking the extreme C-terminus of AEBP1, was able to bind the resin (results not shown). The placement of this hydrophobic stretch of AEBP1 into the protein interior may serve as an anchor to hold the C-terminus in place and restrict accessibility to the highly charged basic and acidic helices within the C-terminus.

The presence of this C-terminal anchor sequence in both AEBP1-WT and AEBP1(123–748) might be the reason for the lack of DNA binding observed for these two proteins, whereas AEBP1(27–730) and AEBP1–ΔSTP both lack this extreme C-terminal sequence and are able to interact strongly with DNA (Fig. 6). It is possible that interaction of AEBP1 with another protein may cause conformational changes that would allow the putative basic helix-containing DNA-binding domain of AEBP1 to be available for DNA binding in vivo. However, both in vitro EMSA and in vivo luciferase assays in CHO cells suggest that WT AEBP1 does not interact with DNA. A lack of DNA binding ability is consistent with a recent report describing the role of ACLP in transdifferentiation of preadipocytes into smooth musclelike cells,[20] in which the authors state that they were unable to detect any binding of AEBP1 to the AE-1 element using a version of AEBP1 lacking the 30 N-terminal amino acids.

AEBP1 Interacts with Ca^{2+}/Calmodulin

AEBP1 was identified as the p85 component of a Ca^{2+}/calmodulin-dependent histone H3 arginine kinase complex through tryptic peptide fingerprinting of p85 by MALDI-TOF mass spectrometry (B. Wakim, personal communication). p85 is the major component of this arginine kinase complex as well as the only component of this complex capable of interacting with Ca^{2+}/calmodulin.[55] The ability of AEBP1 to interact with Ca^{2+}/calmodulin was analyzed and characterized through the use of AEBP1 WT and deletion derivatives in calmodulin–agarose pull-down assays (Fig. 7). Full-length AEBP1 as well as the AEBP1ΔN and AEBP1ΔSTP derivatives showed strong binding to Ca^{2+}/calmodulin that could be abolished by inclusion of EGTA. AEBP1 CP and ΔC both lack the entire C-terminus and both showed only weak association with Ca^{2+}/calmodulin that was not sensitive to the presence of EGTA indicating that it was a nonspecific association.

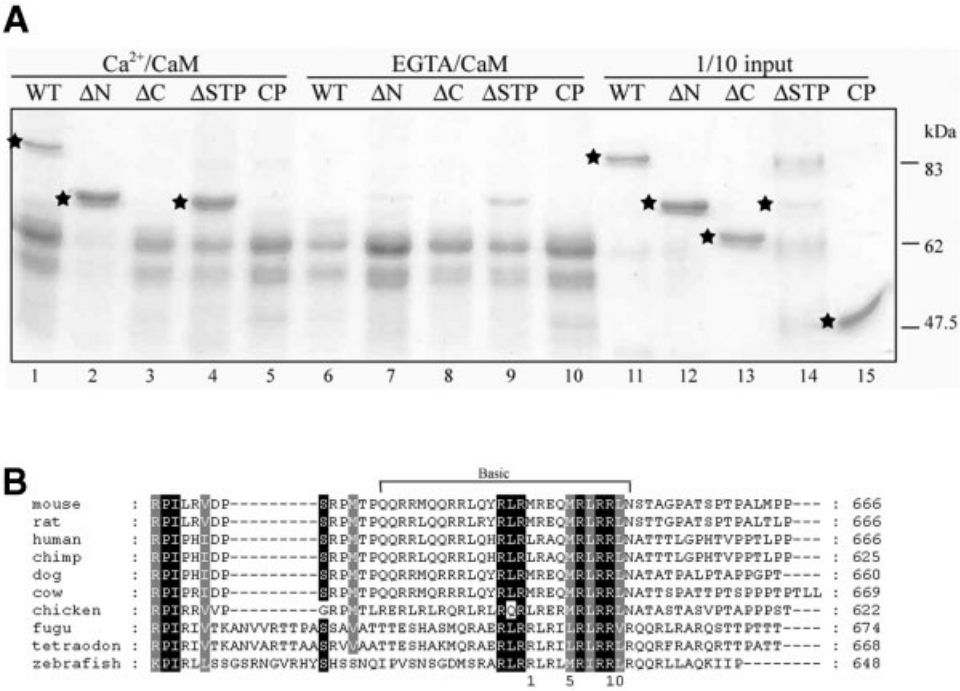


Fig. 7. AEBP1 interacts with Ca^{2+}/calmodulin through its basic region. (**A**) Calmodulin–sepharose was used to pull down AEBP1 and indicated truncation forms of AEBP1 in the presence of Ca^{2+} (lane 1–5) or EGTA (lanes 6–10). Proteins pulled down were separated by SDS-PAGE and stained with GelCode Blue. A star indicates strongly bound AEBP1 as well as input AEBP1 (lanes 11–15). Protein markers were loaded in lane 14 along with AEBP1ΔSTP protein. (**B**) The basic helix of AEBP1 contains the sequence requirements for Ca^{2+}/calmodulin interaction. An alignment of AEBP1 orthologues indicates strong conservation within a portion of the basic domain at the C-terminus of AEBP1. Sequence identity is indicated by black shading, while chemical similarity is indicated by grey shading. The bottom row indicates the residues composing a potential 1–5–10 sequence of hydrophobic amino acids within the basic helix known to be involved in Ca^{2+}/calmodulin interaction by some proteins.

However, AEBP1 ΔSTP, which contains the basic region of the C-terminus, was able to interact with Ca^{2+}/calmodulin very strongly, especially considering that the amount of AEBP1 ΔSTP input into this pulldown was much less than that used in Ca^{2+}/calmodulin pulldowns of AEBP1 WT and other AEBP1 truncations. Multiple lower molecular weight bands were observed in all pulldown lanes following protein staining. These proteins appear to arise from the calmodulin–agarose beads, either calmodulin itself or other copurified proteins. The ability of AEBP1ΔN and AEBP1ΔSTP to interact with Ca^{2+}/calmodulin suggests that neither the DLD domain nor the STP and acidic regions within the C-terminus of AEBP1 are necessary for interaction with Ca^{2+}/calmodulin. The lack of a strong association of Ca^{2+}/calmodulin with AEBP1ΔC suggests that the C-terminus of AEBP1 is necessary for interaction. The only part of the C-terminus that is present in AEBP1ΔSTP and lacking in AEBP1ΔC is the basic region, suggesting that this region is necessary for interaction of AEBP1 with Ca^{2+}/calmodulin. This is the same region implicated in DNA binding by AEBP1. However, while WT AEBP1 is not able to interact with DNA, it is able to interact efficiently with Ca^{2+}/calmodulin.

An interaction of this basic helix with Ca^{2+}/calmodulin is consistent with the current knowledge regarding peptides that interact with Ca^{2+}/calmodulin. While these interacting peptides share no sequence homology, they all potentially fold into an amphipathic α-helix, with large hydrophobic residues either at positions 1–5–10 or 1–8–14.[56] The basic region of AEBP1 can potentially fall into the 1–5–10 class of calmodulin-interacting proteins, as residues M640, M644, and L649 of mouse AEBP1 are all large hydrophobic residues [Fig. 7(b)]. This area of the basic region of AEBP1 is highly conserved across all species. The hydrophobic amino acid in position 1 is not conserved in the fish AEBP1 sequences identified, suggesting that these may not interact with Ca^{2+}/calmodulin or may interact in a different manner. On the other hand, fish AEBP1 may not have the same function as AEBP1 in other vertebrates. Both Fugu and Tetraodon AEBP1 lack a methionine initiation codon in the position where AEBP1 translation is known to initiate from in mice, suggesting that only ACLP may be produced in fish (see Supplementary Material).

CONCLUSIONS

AEBP1 has been previously characterized as a DNA binding transcriptional repressor with carboxypeptidase activity. This work clarifies these functions through structural modeling, bioinformatics, and functional assays. Structural modeling of the discoidin and CP domains of AEBP1 has predicted possible functions for these domains,

much of which has yet to be investigated. The discoidin domain model suggests a role in membrane interaction similar to the C2 domains of coagulation factors V and VIII. The CP domain model suggests a possible function in N-terminal basic residue interaction. As protein translation initiates with a methionine, this interaction would necessarily follow cleavage by an endopeptidase or following methionine N-terminal excision. A lack of conservation of the possible catalytic residues N374 and Y485 raises doubt regarding a role in catalysis. Other laboratories have not been able to reproduce the low CP activity previously reported for AEBP1, altogether suggesting that any activity previously exhibited by this domain may not be physiologically relevant or may have simply been background levels. Finally, a basic region within the C-terminus of AEBP1, predicted to form an α-helix, is shown to be the mediator of DNA binding and Ca^{2+}/calmodulin interaction. However, wild-type AEBP1 and versions of AEBP1 containing the extreme C-terminal hydrophobic region are unable to bind DNA in vitro. Transcriptional assays also indicate that DNA binding is not necessary for transcriptional repression in CHO cells. These results greatly modify our understanding of AEBP1, which may function as a transcriptional corepressor using its CP domain in an interaction function only. Molecular modeling and sequence analysis presented here open the doors to many potential investigations into the structure and function of ACLP/AEBP1, as well as the related proteins CPX-1 and CPX-2.

REFERENCES

1. Enard W, Khaitovich P, Klose J, et al. 2002. Intra- and interspecific variation in primate gene expression patterns. Science 296: 340–343.
2. Pennisi E. 2004. Searching for the genome's second code. Science 306:632–635.
3. Hahn S. 2004. Structure and mechanism of the RNA polymerase II transcription machinery. Nat Struct Mol Biol 11:394–403.
4. Babu MM, Luscombe NM, Aravind L, et al. 2004. Structure and evolution of transcriptional regulatory networks. Curr Opin Struct Biol 14:283–291.
5. Flaus A, Owen-Hughes T. 2004. Mechanisms for ATP-dependent chromatin remodelling: farewell to the tuna-can octamer? Curr Opin Genet Dev 14:165–173.
6. Kamakaka RT, Biggins S. 2005. Histone variants: deviants? Genes Dev 19:295–310.
7. de la Cruz X, Lois S, Sanchez-Molina S, et al. 2005. Do protein motifs read the histone code? Bioessays 27:164–175.
8. Peterson CL, Laniel MA. 2004. Histones and histone modifications. Curr Biol 14:R546–R551.
9. Wu J, Grunstein M. 2000. 25 years after the nucleosome model: chromatin modifications. Trends Biochem Sci 25:619–623.
10. Vogelauer M, Wu J, Suka N, et al. 2000. Global histone acetylation and deacetylation in yeast. Nature 408:495–498.
11. Zhang Y, Reinberg D. 2001. Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. Genes Dev 15:2343–2360.
12. Nowak SJ, Corces VG. 2004. Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. Trends Genet 20:214–220.
13. Gill G. 2004. SUMO and ubiquitin in the nucleus: different functions, similar mechanisms? Genes Dev 18:2046–2059.
14. Turner BM. 2002. Cellular memory and the histone code. 111:285–291.
15. He GP, Muise A, Li AW, et al. 1995. A eukaryotic transcriptional repressor with carboxypeptidase activity. Nature 378:92–96.
16. Gregoire FM, Smas CM, Sul HS. 1998. Understanding adipocyte differentiation. Physiol Rev 78:783–809.
17. Kim SW, Muise AM, Lyons PJ, et al. 2001. Regulation of adipogenesis by a transcriptional repressor that modulates MAPK activation. J Biol Chem 276:10199–10206.
18. Ohno I, Hashimoto J, Shimizu K, et al. 1996. A cDNA cloning of human AEBP1 from primary cultured osteoblasts and its expression in a differentiating osteoblastic cell line. Biochem Biophys Res Commun 228:411–414.
19. Park JG, Muise A, He GP, et al. 1999. Transcriptional regulation by the gamma5 subunit of a heterotrimeric G protein during adipogenesis. EMBO J 18:4004–4012.
20. Abderrahim-Ferkoune A, Bezy O, Astri-Roques S, et al. 2004. Transdifferentiation of preadipose cells into smooth muscle-like cells: role of aortic carboxypeptidase-like protein. Exp Cell Res 293:219–228.
21. Layne MD, Endege WO, Jain MK, et al. 1998. Aortic carboxypeptidase-like protein, a novel protein with discoidin and carboxypeptidase-like domains, is up-regulated during vascular smooth muscle cell differentiation. J Biol Chem 273:15654–15660.
22. Layne MD, Yet SF, Maemura K, et al. 2001. Impaired abdominal wall development and deficient wound healing in mice lacking aortic carboxypeptidase-like protein. Mol Cell Biol 21:5256–5261.
23. Lei Y, Xin X, Morgan D, et al. 1999. Identification of mouse CPX-1, a novel member of the metallocarboxypeptidase gene family with highest similarity to CPX-2. DNA Cell Biol 18:175–185.
24. Xin X, Day R, Dong W, et al. 1998. Identification of mouse CPX-2, a novel member of the metallocarboxypeptidase gene family: cDNA cloning, mRNA distribution, and protein expression and characterization. DNA Cell Biol 17:897–909.
25. Altschul SF, Madden TL, Schaffer AA, et al. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402.
26. Fiser A, Do RK, Sali A. 2000. Modeling of loops in protein structures. Protein Sci 9:1753–1773.
27. Sali A, Blundell TL. 1993. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779–815.
28. Sippl MJ. 1993. Recognition of errors in three-dimensional structures of proteins. Proteins 17:355–362.
29. Laskowski RA, MacArthur MW, Moss DS, et al. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst 26:283–291.
30. Canutescu AA, Shelenkov AA, Dunbrack RL Jr. 2003. A graph-theory algorithm for rapid protein side-chain prediction. Protein Sci 12:2001–2014.
31. Kraulis PJ. 1991. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J Appl Cryst 24:946–950.
32. Merritt EA, Bacon DJ. 1997. Raster3D photorealistic molecular graphics. Meth Enzymol 277:505–524.
33. Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680.
34. Nicholas KB, Nicholas HB Jr, Deerfield DW II. 1997. GeneDoc: analysis and visualization of genetic variation. EMBNEWNEWS 4:14.
35. Page RD. 1996. TreeView: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12:357–358.
36. Macedo-Ribeiro S, Bode W, Huber R, et al. 1999. Crystal structures of the membrane-binding C2 domain of human coagulation factor V. Nature 402:434–439.
37. Pratt KP, Shen BW, Takeshima K, et al. 1999. Structure of the C2 domain of human factor VIII at 1.5 A resolution. Nature 402:439–442.
38. Gomis-Ruth FX, Companys V, Qian Y, et al. 1999. Crystal structure of avian carboxypeptidase D domain II: a prototype for the regulatory metallocarboxypeptidase subfamily. EMBO J 18: 5817–5826.
39. Novikova EG, Eng FJ, Yan L, et al. 1999. Characterization of the enzymatic properties of the first and second domains of metallocarboxypeptidase D. J Biol Chem 274:28887–28892.
40. Aloy P, Companys V, Vendrell J, et al. 2001. The crystal structure of the inhibitor-complexed carboxypeptidase D domain II and the modeling of regulatory carboxypeptidases. J Biol Chem 276:16177–16184.
41. Cumming RC, Andon NL, Haynes PA, et al. 2004. Protein disulfide bond formation in the cytoplasm during oxidative stress. J Biol Chem 279:21749–21758.

42. Cleasby A, Wonacott A, Skarzynski T, et al. 1996. The x-ray crystal structure of phosphomannose isomerase from *Candida albicans* at 1.7 angstrom resolution. Nat Struct Biol 3:470–479.
43. Roach PL, Clifton IJ, Fulop V, et al. 1995. Crystal structure of isopenicillin N synthase is the first from a new structural family of enzymes. Nature 375:700–704.
44. Sami M, Brown TJ, Roach PL, et al. 1997. Glutamine-330 is not essential for activity in isopenicillin N synthase from Aspergillus nidulans. FEBS Lett 405:191–194.
45. Lesburg CA, Huang C, Christianson DW, et al. 1997. Histidine → carboxamide ligand substitutions in the zinc binding site of carbonic anhydrase II alter metal coordination geometry but retain catalytic activity. Biochemistry 36:15780–15791.
46. Yiallouros I, Grosse Berkhoff E, Stocker W. 2000. The roles of Glu93 and Tyr149 in astacin-like zinc peptidases. FEBS Lett 484:224–228.
47. Blomster M, Wetterholm A, Mueller MJ, et al. 1995. Evidence for a catalytic role of tyrosine 383 in the peptidase reaction of leukotriene A4 hydrolase. Eur J Biochem 231:528–534.
48. Thunnissen MM, Andersson B, Samuelsson B, et al. 2002. Crystal structures of leukotriene A4 hydrolase in complex with captopril and two competitive tight-binding inhibitors. FASEB J 16:1648–1650.
49. Muise AM, Ro HS. 1999. Enzymic characterization of a novel member of the regulatory B-like carboxypeptidase with transcriptional repression function: stimulation of enzymic activity by its target DNA. Biochem J 343:341–345.
50. Qian Y, Varlamov O, Fricker LD. 1999. Glu300 of rat carboxypeptidase E is essential for enzymatic activity but not substrate binding or routing to the regulated secretory pathway. J Biol Chem 274:11582–11586.
51. Song L, Fricker LD. 1997. Cloning and expression of human carboxypeptidase Z, a novel metallocarboxypeptidase. J Biol Chem 272:10543–10550.
52. Arminjon F, Guinand M, Vacheron MJ, 1977. Specificity profiles of the membrane-bound gamma-D-glutamyl-(L)meso- diaminopimelateendopeptidase and LD-carboxypeptidase from *Bacillus sphaericus* 9602. Eur J Biochem 73:557–565.
53. Garnier M, Vacheron MJ, Guinand M, et al. 1985. Purification and partial characterization of the extracellular gamma-D- glutamyl-(L)meso-diaminopimelate endopeptidase I, from Bacillus sphaericus NCTC 9602. Eur J Biochem 148:539–543.
54. Hourdou ML, Guinand M, Vacheron MJ, et al. 1993. Characterization of the sporulation-related gamma-D-glutamyl-(L)meso- diaminopimelic-acid-hydrolysing peptidase I of *Bacillus sphaericus* NCTC 9602 as a member of the metallo(zinc) carboxypeptidase A family. Modular design of the protein. Biochem J 292:563–570.
55. Wakim BT, Grutkoski PS, Vaughan AT, et al. 1995. Stimulation of a Ca(2+)-calmodulin-activated histone 3 arginine kinase in quiescent rat heart endothelial cells compared to actively dividing cells. J Biol Chem 270:23155–23158.
56. Vetter SW, Leclerc E. 2003. Novel aspects of calmodulin target recognition and activation. Eur J Biochem 270:404–414.
57. Lyons PJ, Muise AM, Ro HS. 2005. MAPK modulates the DNA binding of adipocyte enhancer-binding protein 1. Biochemistry 44:926–931.