cC/EPB, a chicken transcription factor of the leucine-zipper C/EBP family

Cor F.Calkhoven, Geert AB* and Jan Wijnholds
Department of Biochemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

Submitted June 30, 1992 EMBL accession no. X66844

The gene encoding the chicken homologue of the rat CC-AAT/Enhancer Binding protein C/EBPα (1, 2) was cloned and sequenced. The strategy consisted of the screening of a chicken liver cDNA library with a rat C/EBP clone yielding a partial cDNA (3), which in turn served as a probe to screen a chicken genomic DNA library. Of two independent clones, a 1686-bp region covering the C/EBP gene was sequenced. Both sequences are identical and contain a 971-bp open reading frame. A putative TATA box is present at 230 bp 5′ and a putative polyadenylation signal at 285 bp 3′ to the open reading frame.

Comparative sequence analysis shows that the derived 324-amino acid chicken C/EBP sequence is highly similar (68.5%) to the 358-residue C/EBPα sequence of rat (Figure 1). The C-terminal moieties constituting the basic, DNA-binding and leucine zipper, dimerisation domains are virtually identical (94%). The N-terminal moieties are partially conserved (59%); three highly conserved regions, designated I, II and III (Figure 1), can be distinguished and may correspond to individual functional domains. Our conserved region I coincides with the N-terminal *trans*-acting sequence defined in rat C/EBPα. Regions II and III map in a region of which the function is not clear (4, 5). The conserved regions revealed in our investigations may help to further define the functional domains. Chicken C/EBP differs notably from the rat C/EBPα in not having the proline and glycine stretches lying between the conserved regions.

In the 5′ untranslated sequence of chicken C/EBP, we find a small open reading frame for a hypothetical peptide of 5 amino acids (MPGRL) separated by 7 nt from the C/EBP reading frame. A sequence potentially encoding a similar peptide (MPGEL) is present at exactly the same position in the rat C/EBPα gene. The cognate DNA sequence only shows moderate similarity (11/18), suggesting that conservation is imposed at the amino acid level.

As shown earlier by cotransfection experiments with a rat C/EBPα expression vector, C/EBP is a potential factor involved in the activation of the liver-specific, estrogen inducible apoVLDL II gene (6). In view of the sequence differences found, it will be interesting to compare the activities of the homologous cC/EBP in transfection experiments.

REFERENCES

1. Landschulz,W.H., Johnson P.F., Adashi,E.Y., Graves,B.J. and McKnight,S.L. (1988) *Genes Dev.* **2**, 786–800.
2. Cao,Z., Umek,R.M. and McKnight,S.L. (1991) *Genes Dev.* **5**, 1538–1552.
3. Wijnholds,J. (1991) Thesis, Groningen University.
4. Friedman,A.D. and McKnight,S.L. (1989) *Genes Dev.* **4**, 1416–1426.
5. Pei,D. and Shih,C. (1991) *Mol. Cell. Biol.* **11**, 1480–1487.
6. Beekman,J.M., Wijnholds,J., Schippers,I.J., Pot,W., Gruber,M. and AB,G. (1991) *Nucleic Acids Res.* **19**, 5371–5377.

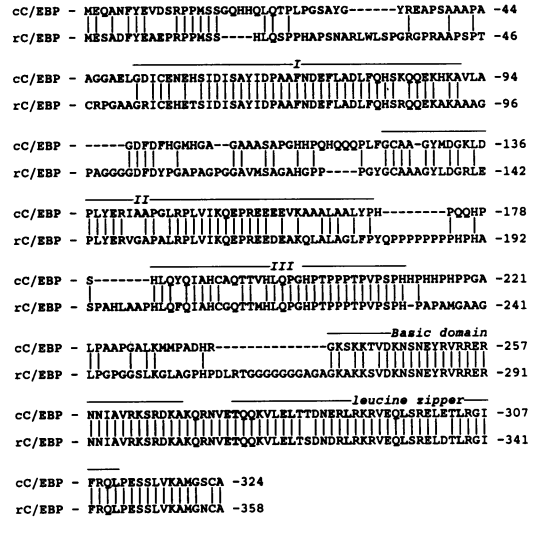

Figure 1. Alignment of the chicken C/EBP (cC/EBP) and rat C/EBPα (rC/EBP) amino acid sequences. Highly conserved domains are indicated: Conserved regions I, II and III; the basic DNA-binding domain and the leucine zipper dimerisation domain. Amino acid identities are indicated by bars. The numbering on the right refers to the position of the amino acids. The partial cDNA sequence (see text) encodes amino acids 154 through 313.

* To whom correspondence should be addressed