Isolation and characterization of cDNAs encoding the sea urchin (*Strongylocentrotus purpuratus*) homologue of the CCAAT binding protein NF-Y A subunit

Zhe Li, Sirinivas R. Kalasapudi and Geoffrey Childs*

Department of Molecular Genetics, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, USA

Received July 29, 1993; Accepted July 30, 1993

GenBank accession no. L19224

We have isolated the sea urchin homologue of the mammalian transcription factor NF-Y A subunit (CP-1, CBP) (1, 2). The cloned murine NF-Y A cDNA (3) was used as a template for PCR reactions to amplify the region between amino acids 237–275. This region of the protein is highly conserved in species ranging from yeast to humans (4) and the subunits are functionally interchangeable (5). Low stringency hybridization conditions were used to screen a sea urchin *Strongylocentrotus purpuratus* cDNA library prepared from 4 cell embryos (4 hr post-fertilization). The nucleotide sequence of the cDNA contains an open reading frame of 400 amino acids. The sea urchin protein is larger than the yeast Hap2 protein, 265 amino acids (6), or the murine, 317 and 346 amino acids (3) and human 347 (3) or 257 (7) amino acids, NF-Y A subunits.

Comparison of the amino acid sequences from the different species confirms the high degree of sequence similarity within the Hap2 core domain (6; Figure 1). Within a stretch of 69 amino acid residues there is a 93% sequence identity between the sea urchin and mammalian polypeptides. Among the five differences in this region, only one, a valine (sea urchin) to lysine (mammalian) permutation at position 355 of the sea urchin protein is not a conservative change. This position is variable in different species and is a proline in the yeast Hap2 protein (6). The 69 amino acids of sequence similarity between the sea urchin and mammalian protein is 13 residues longer than the region conserved between the yeast Hap2 and mammalian NF-Y A proteins. The extra highly conserved amino acids are at the N-terminal side of this domain that includes the region responsible for dimerization and DNA binding (4). While sequence comparisons between kingdoms revealed sharp boundaries between the conserved domain and diverged sequences (4), the sea urchin protein has an additional region of lower homology that extends for an additional 50 amino acid residues towards the N-terminus. There is only limited sequence conservation in the remaining regions of the proteins. The most striking additional feature of all the NF-Y A proteins is a glutamine rich region that can act as a transcriptional activation domain. The sea urchin protein has 10 consecutive glutamine residues within the region that is also rich in threonine and serine.

The sea urchin mRNA encoding NF-Y A is a single 3.9 kb maternal polyadenylated transcript (data not shown). The transcript is extremely rare and can barely be detected using 15 µg of maternal (egg) poly(A)⁺ mRNA. The transcript is also

detected in morula and early blastula stage embryonic mRNA but not at later stages of development. We estimate that there are 2000 or fewer transcripts stored in the egg and the total number of transcripts per embryo is relatively constant until the 128-cell blastula stage of development. We have not looked for NF-Y A transcripts in different adult tissues but a protein with the characteristics of NF-Y has been identified in sea urchin testis (8).

ACKNOWLEDGMENTS

We would like to thank Drs. Diane Mathis for kindly providing us with mouse cDNAs encoding NF-Y and Eric Davidson for the cDNA library. This work was supported by NIH grant GM30333.

REFERENCES

1. Dorn, A. *et al.* (1987) *Cell* **50**, 863–872.
2. Chodosh, L.A. *et al.* (1988) *Cell* **53**, 11–24.
3. van Huijsduijn, R.H. *et al.* (1990) *EMBO J.* **9**, 3119–3127.
4. Li, X.Y. *et al.* (1992) *Nucleic Acids Res.* **20**, 1087–1091.
5. Chodosh, L.A. *et al.* (1988) *Cell* **53**, 25–35.
6. Pinkham, J.L. *et al.* (1987) *Mol. Cell Biol.* **7**, 578–585.
7. Becker, D.M. *et al.* (1991) *Proc. Nat. Acad. Sci. USA* **88**, 1968–1972.
8. Barberis, A. *et al.* (1987) *Cell* **50**, 347–359.

Figure 1. Alignment of the amino-acid sequence of sea urchin and murine NF-Y A proteins. Vertical lines indicate identical residues and dots conservative changes. Sequences were analyzed using the GCG package. The sea urchin sequence is on the top line and the mouse sequence (3) below.

* To whom correspondence should be addressed