Full length cDNA sequence encoding a nuclease-sensitive element DNA binding protein

Rukmini Kolluri and Alan J.Kinniburgh*

Department of Human Genetics, Roswell Park Cancer Institute, Elm and Carlton Streets, Buffalo, NY 14263, USA

Submitted July 9, 1991

GenBank accession no. M38356

The *c-myc* gene contains a positive *cis*-acting promoter element (the nuclease sensitive element or NSE) that we have previously shown binds nuclear factors (1). *In vitro*, this DNA region can assume an H-DNA conformation (intramolecular triplex/single-strand structure) and is sensitive to single-strand nuclease (1, 2). An oligonucleotide containing this element was used to screen a HeLa cell cDNA expression library. A cDNA clone encoding a 322 a.a. DNA binding protein was isolated, which we term NSEP-1. NSEP-1 binds both duplex and the C-rich strand of the *c-myc* NSE DNA in a sequence specific manner. NSEP-1 has been reported previously as a relatively uncharacterised protein (dbpB) that binds to an EGFR promoter element, a sequence highly related to the *c-myc* NSE (3). We report this sequence because the previously reported sequence is inconsistent with our results (Figure 1). Various deleted recombinant NSEP-1 proteins that we have made are consistent with our sequence whereas the previously reported amino acid sequences would lead to nonsense peptides. Therefore the previously reported sequence is either incorrect or is highly related to but genetically distinct from NSEP-1. The deduced amino acid sequence shows four putative DNA binding domains; three are rich in basic amino acids. Two of these basic regions, basic-1 and basic-2 form the double-strand DNA binding domain of the NSEP-1 protein. In addition, NSEP-1 also contains a Pro/Ser/Thr rich domain and a separate domain rich in Asp/Glu/Gln, both reminiscent of other factors' activation domains (Figure 2). A unique feature of the NSEP-1 protein is the presence of an octapeptide single-strand DNA binding motif which bears weak homology to the ribonucleoprotein consensus sequence (RNP-CS-1); a structural feature common to RNA binding proteins (our unpublished data). Since protein segments containing primarily this region are able to bind single strand polynucleotides, this region could confer the single-strand binding specificity of NSEP-1.

ACKNOWLEDGEMENTS

This work was supported by a grant from the NIH, CA43661, and a Cardiovascular Discovery Award from Glaxo, Inc. We thank Nancy Frame for secretarial assistance.

REFERENCES

1. Davis, T.L., Firulli, A.B. and Kinniburgh, A.J. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9682–9686.
2. Kinniburgh, A.J. (1989) *Nucl. Acids. Res.* **17**, 7771–7778.
3. Sakura, H., Maekwa, T., Imamoto, F., Yasuda, K. and Ishii, S. (1988) *Gene* **73**, 499–507.

	680v	690v	710v	720v	730v	740v	860v	870v
NSEP1	CTACCGG-AGCGGGCAGGGAGCGGTG	GGGGC-TACACATCGCGCGCGCTGGCGGGG	GACAA	TAAAGGGGGGAAGAATAAC				
M24070	CTAC CG AGCGGGCAGGGAGCGGTG	GGGGC TCACATCGCGCGCGCTG	GC G G GACAA TAAAG	ANGAATAAC				
dbpB	210^	220^	230^	0^	250^	260^	270^	400^
	1130v	1140v	1180v	1190v	1220v	1230v	1550v	1560v
NSEP1	CCCGAAG-CCAGGCCCAACA	TACTACATGCGGAGACC-TAT	CAACCCCT	GTGCAGGGAG	GGGGCGGGGCTGACT			
M24070	CCCGAAG CCAGGCCCAACA	TACTACATGCGGAGACC TAT	CAACCCCT	GTGCAGGGAG	GGG CGGGGCTGACT			
dbpB	660^	670^	710^	720^	750^	760^	-	1090^

Figure 1. Deviations from the published sequence of dbpB.

Figure 2. Structural features of NSEP-1.

* To whom correspondence should be addressed