Vol. 164, No. 1, 1989 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
October 16, 1989 Pages 149-155

THE SAME NUCLEAR PROTEINS BIND THE PROXIMAL CACCC BOX OF THE HUMAN β-GLOBIN PROMOTER AND A SIMILAR SEQUENCE IN THE ENHANCER

Barbara Giglioni•, Paola Comi*•, Antonella Ronchi**, Roberto Mantovani**
and Sergio Ottolenghi**•

•Centro per lo Studio sulla Patologia Cellulare, C.N.R., Milano, Italy

*Istituto di Patologia Generale, Milano, Italy

**Dipartimento di Genetica e di Biologia dei Microrganismi, Milano, Italy

Received August 4, 1989

Summary: Using *in vitro* assays, we show that nuclear proteins related to the Sp1 and GT-I factors bind to a CACCC box sequence in the human β-globin enhancer, adjacent to binding sites for the erythroid-specific factor NFE1 and the ubiquitous factor CP1. The same proteins are known to bind to the proximal, but not to the distal, CACCC box in the human β-globin promoter. A C G mutation in the promoter CACCC box, known to cause β-thalassemia, greatly decreases protein binding to the CACCC box; the same effect is obtained when this mutation is introduced into the enhancer CACCC box.

Sequences within enhancer elements greatly increase the activity of promoters, producing multiplicative, rather than simply additive, effects (1). According to one often quoted model (2), distant sequences in enhancer and promoter elements are brought into close physical proximity by proteins (possibly dimers) interacting with sequences present in both elements. In the human and mouse β-globin promoters, several conserved DNA motifs capable of binding potential transcriptional factors, have been identified (3-5), and the functional role of some of them (CACCC, CCAAT, NFE1) has been documented in transfection experiments (3,6,7). Moreover, β-thalassemic mutations (8-9), identified in several populations, directly point to the CACCC box as a major transcription regulatory element of the β-globin gene promoter. The correlation between presence of an intact CACCC box and activity of the β-globin promoter, is further stressed by the observation that a β-thalassemic mutation (-87 C►G), affecting this element and known to greatly decrease *in vivo* β-globin transcription, almost abolishes the binding of nuclear proteins to the mutated promoter; significantly, a CACCC box-like sequence lying distal to the mutated CACCC box is unable to maintain the normal activity of the promoter, and binds little, if any, nuclear protein (5).

A region of the β-globin enhancer includes a sequence with great similarity to the promoter CACCC sequence shown to be able to bind proteins; this sequence is

0006-291X/89 $1.50

therefore a candidate for an enhancer element capable of interacting (via protein-binding) with the promoter region. Here we show that the enhancer CACCC box region does indeed bind the same proteins that are able to interact with the proximal CACCC box of the promoter; the identified enhancer CACCC box lies in close proximity to sites capable of binding the erythroid specific protein NFE1 (4,10,11) and the ubiquitous factor CP1 (11).

Materials and Methods

Preparation of nuclear extracts

Nuclear extracts were prepared exactly according to Dignam et al. (12) from exponentially growing K562 cells.

Oligonucleotides

The sequence of the oligonucleotides used in this work is given in Fig. 3D (β-globin enhancer, Sp1 and GT-I) and ref. 5 (β-globin promoter). For Sp1 and GT-I oligonucleotides, see also ref. 13.

Electrophoretic mobility shift assay (14)

One strand of the oligonucleotide to be used was 5'-end-labelled *in vitro* with [^{32}P] using T4 polynucleotide Kinase and annealed with a 5-fold excess of the complementary oligonucleotide in 67 mM Tris HCl pH 7.8, 13 mM $MgCl_2$, 6.7 mM DTT, 1.3 mM spermidine, 1.3 mM EDTA by heating to 85°C for 2 min. followed by slowly lowering the temperature to 4°C, in a 2 hours period. The annealed oligonucleotide was then purified by gel electrophoresis. Binding reactions and electrophoresis were carried out essentially according to ref. 5. Briefly, the standard assay contained (in 20 μl): 0.1-0.2 ng of [^{32}P] labelled oligonucleotide, 3-6 μg of nuclear protein, 3 μg of poly (dI-dC) in a buffer containing 4 mM spermidine, 50 mM NaCl, 1 mM EDTA, 10 mM Tris HCl pH 7.9. 1 mM DTT, 0.5 mM PMSF. Unlabelled competitor DNAs were as specified in Figure legends. Incubation was at 20°C for 30'. Gel electrophoresis was in 5% acrylamide gels in 50 mM Tris borate pH 8.2.

DMS interference (15)

Oligonucleotides 5'-end labelled at only one end, were partially methylated with DMS (1 μl/20μl reaction containing 50 mM Na Cacodylate pH 8, 10 mM $MgCl_2$, 0.1 mM EDTA; incubation was at 20°C for 11 min.). After stopping the reaction with 200 mM Tris Acetate, 0.2 M β-mercaptoethanol, 0.2 mM EDTA, the fragment was ethanol precipitated, incubated with nuclear extracts and run on polyacrylamide gels. Bands were eluted, treated with 1M piperidine for 30 min. at 90°C, lyophilized and analysed on sequencing gels.

Results

Figure 1 shows the nucleotide sequence of the region adjacent to the NFE1 binding site B of the human β-globin enhancer (10,11); a sequence closely resembling the CACCC box present in the β-globin promoter is present immediately 3' to the NFE1 binding site (this sequence was not represented in the oligonucleotide used, in ref. 11). Xiao et al. (13) previously reported that the mouse β-globin promoter CACCC box binds *in vitro* to two different proteins: the ubiquitous factor Sp1 and the partially tissue-specific GT-I; to investigate the possibility that the human β-globin enhancer (and promoter) CACCC box could bind the same proteins, we compared the gel-shift patterns obtained (with K562 nuclear extracts) with oligonucleotides from the two regions (Fig. 1) and with Sp1 and GT-I consensus sequences (13).

The Sp1 oligonucleotide gives four bands, labelled 1 to 4 in Fig. 2, lane 1, (bands 1 and 2 are poorly resolved and visible only in underexposed autoradiograps); GT-I

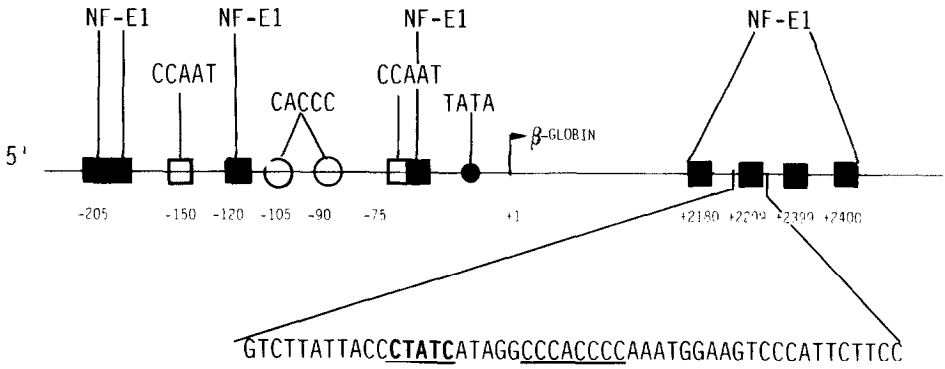


Fig. 1 Nuclear protein binding sites in the β-globin promoter and enhancer. The NFE-1 binding sequence and the putative CACCC box in the β-globin enhancer site B (10,11) are shown below the diagram.

gives the same four bands, but bands 2 and 3 are much fainter (lane 2). Both the enhancer (lane 3) and promoter oligonucleotides (lane 10) give a very similar pattern to Sp1; however, with the enhancer oligonucleotide band 2 is prevalent over band 1, while the reverse is true for the promoter oligonucleotide.

Additional bands are generated by the enhancer oligonucleotide; a slow band (asterisk) is due to interaction with a CCAAT-box binding protein CP1 (11), as demonstrated by competition (not shown) with excess unlabelled γ-globin promoter CCAAT-box containing oligonucleotide (10); a faint band due to interaction with NFE1

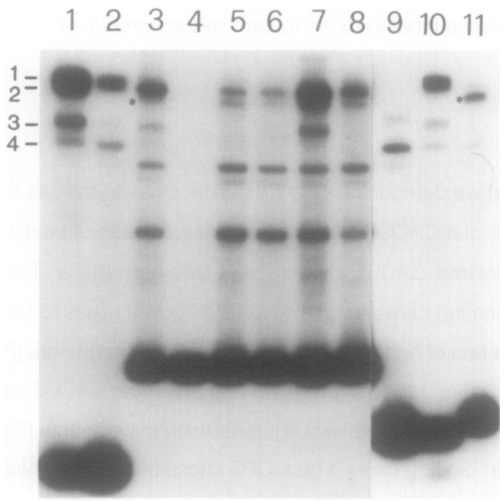


Fig. 2 Binding of CACCC containing oligonucleotides from β-globin enhancer and promoter regions to K562 nuclear extracts. Lane 1: control Sp1 oligonucleotide; lane 2: GT-I; lane 3: enhancer oligonucleotide; lanes 4-8: enhancer oligonucleotide competed by unlabelled enhancer (200-fold excess), Sp1, GT-I, unrelated oligonucleotide (mouse α1-globin) and β-globin promoter oligonucleotide (all at a 100-fold excess) respectively; lanes 9-11: control NFE-1, β-globin promoter, γ-CCAAT box containing oligonucleotides. The asterisk indicates the band due to interaction with CP1 (11).

(10,11) is almost superimposed with band 4 and is seen clearly using truncated oligonucleotides. Other bands have not been further investigated.

All bands disappear upon competition with excess unlabelled enhancer oligonucleotide (Fig. 2, lane 4); however, competition with unlabelled Sp1, GT-I or β-globin promoter oligonucleotides, greatly decreases the intensities of bands 1 to 4, without interfering with other bands (Fig. 2, lanes 5,6,8). Several control oligonucleotides carrying sequences unrelated to the CACCC box (for example the NFE1 binding α1-mouse globin oligonucleotide of ref. 10) do not compete any of the four complexes (data not shown and Fig. 2, lane 7). These data suggest that the CACCC box regions of the human β-globin enhancer (and promoter) might bind these proteins. Therefore, we analysed the DMS interference pattern of the major bands generated by interaction of nuclear proteins with the enhancer oligonucleotide (Figure 3). This analysis shows that binding is affected by methylation of the G-rich sequence complementary to the CACCC box (bottom strand, Fig. 3A, lanes 1-5) and of the two guanines immediately 5' to the CACCC box (top strand, Fig. 3B). A similar result had been previously obtained with the proximal (but not the distal) CACCC box of the human β-globin promoter (5). The sequence demonstrated to be important for binding is very similar to that essential for Sp1 (Fig. 3C, lanes 1-4) and GT-I binding (13); see diagram in Fig. 3D.

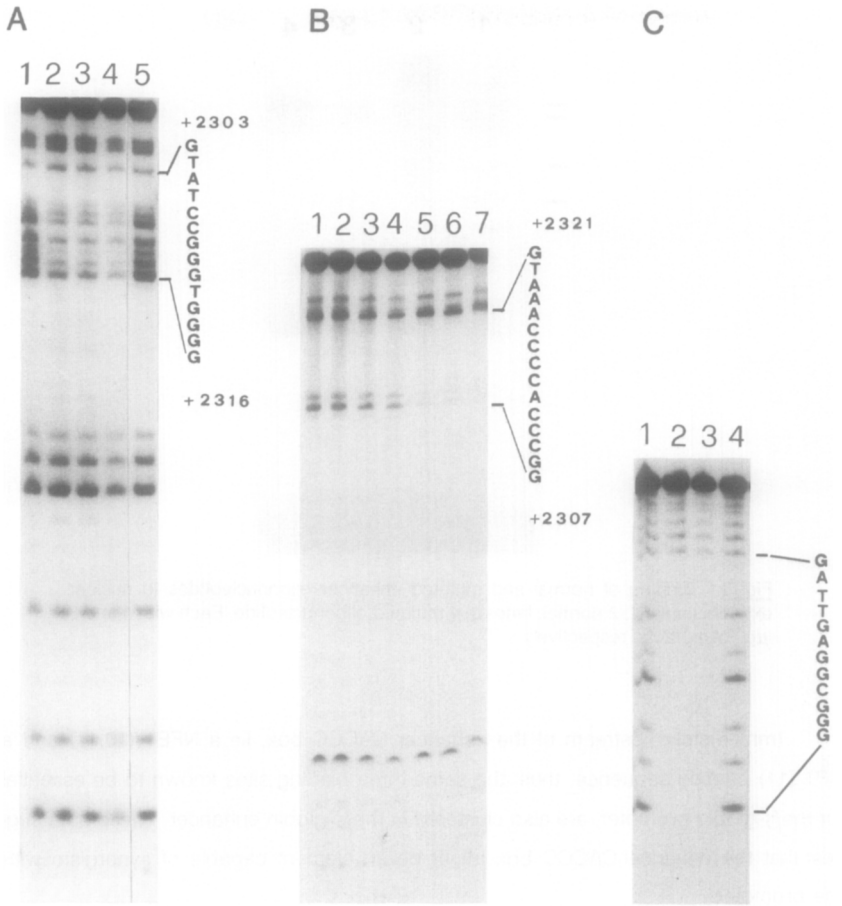
One β-thalassemic mutation (-87 C→G), severely affecting β-globin transcription, disrupts the proximal CACCC box of the human β-globin promoter (8,9) and greatly decreases nuclear protein binding to the CACCC box (5); we have introduced a similar mutation in the β-globin enhancer CACCC box, and we similarly demonstrate a substantially decreased binding to the mutant sequence (Fig. 4).

Discussion

In this report we demonstrate the existence, close to the site B (11) of the human β-globin enhancer, of a CACCC-box like sequence capable of binding *in vitro* the same proteins as the proximal CACCC box of the β-globin promoter. The functional identity (in binding experiments) between the two CACCC boxes is based on:

- similarity of migration of bands generated in gel shift experiments (Fig. 2);
- competition for binding between promoter and enhancer CACCC sequences (Fig. 2);
- similarity of DMS effects in methylation interference experiments (Fig. 3);
- greatly decreased binding caused by a C→G substitution in the third C of the CACCC sequence of both the promoter and enhancer regions (Fig. 4).

It is of interest that the distal CACCC box of the β-globin promoter binds weakly, if at all, to the Sp1 and GT-I proteins (5); this CACCC motif is preceded by a T, while the proximal and the enhancer motifs have either a C or an A at this position, like GT-I and Sp1; similarly, the mouse CACCC motif, capable of binding the same proteins, has A at this position (Table 1).



D

β-ENHANCER

```
                              G
5'- GTCTTATTACCCTATCATAGGCCCACCCCAAATGGAAGTCCCATTCTTCC
    CAGAATAATGGGATAGTATCCGGGTGGGGTTTACCTTCAGGGTAAGAAGG

β-promoter
5'- TGTGAAGCCACACCCTAGGGTT
    ACACCTCGGTGTGGGATCCCAA

Sp 1
5'- GCCCCTAACTCCGCCCAGTTC
    CGGGGATTGAGGCGGGTCAAG

GT-I
5'- GTCAGTTAGGGTGTGGAAAGTC
    CAGTCAATCCCACACCTTTCAG
```

Fig. 3 DMS interference experiment with the β-globin enhancer oligonucleotide bottom (A) and top (B) strands, and with control Sp1 oligonucleotide, bottom strand (C). In A, lanes 1 and 5: unbound oligonucleotide; lanes 2-3: bands 1+2; lane 4: band 4. In B, lanes 1-4 different dilutions of unbound oligonucleotide, lanes 5,6: bands 1+2; lane 7: band 3. In C, lanes 1 and 4: unbound oligonucleotide; lanes 2,3: bands 1+2 from Sp1 oligonucleotide. The interpretation of the data is shown in D; data for GT-I are from ref. 13; for β-globin promoter from ref. 5. ■ ■ indicate G-residues whose methylation affects binding.

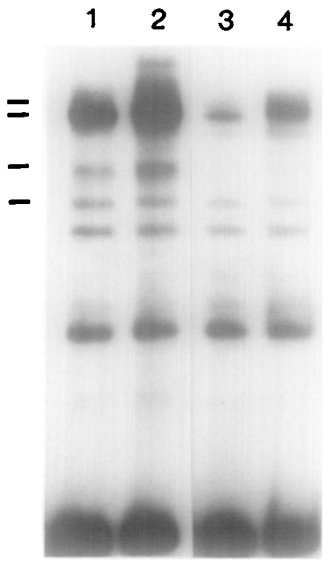


Fig. 4 Binding of normal and mutated enhancer oligonucleotides to nuclear extracts (lanes 1, 2 normal; lanes 3, 4 mutated oligonucleotide. Each with 4 and 8 μg of extracts, respectively).

Immediately upstream of the enhancer CACCC box, lie a NFE1 (10,11) and a CP1 (11) binding sequence; thus, the same three binding sites known to be essential for the β-globin promoter, are also clustered in the β-globin enhancer. These data suggest that the enhancer CACCC box might be an element capable of synergism with the promoter.

Acknowledgment

This work was partially supported by C.N.R. grant 87.00866.51 Progetto Finalizzato Ingegneria Genetica e Basi Molecolari delle Malattie Ereditarie to S.O.

Table 1 Comparison of GT-I and Sp1 binding motifs to CACCC box regions of β-globin promoters and enhancer

Human β-globin Promoter (distal)	5'- agaCCTCACCCtgtgg	top strand
Human β-globin Promoter (proximal)	5'- gagCCACACCCtgtgg	"
Human β-globin Enhancer	5'- tagGCCCACCCcaaat	"
Mouse β-globin Promoter	5'- gagCCACACCCtggta	"
GT-I	5'- tttCCACACCCtaact	bottom
Sp1	5'- taaCTCCGCCCagttc	top

References

1. Muller, M.M., Gerster, T., and Schaffner, W. (1988) Eur. J. Biochem. 176, 485-495.
2. Ptashne, M. (1986) Nature 322, 697-701.
3. de Boer, E., Antoniou, M., Mignotte, V., Wall, L., and Grosveld, F. (1988) EMBO J. 7, 4203-4212.
4. Plumb, M., Frampton, J., Wainwright, H., Walker, M., Macleod, K., Goodvin, G., and Harrison, P. (1989) Nucleic Acids Res. 17, 73-93.
5. Mantovani, R., Malgaretti, N., Nicolis, S., Giglioni, B., Comi, P., Cappellini, N., Bertero, M.T., Caligaris-Cappio, F., and Ottolenghi, S. (1988) Nucleic Acids Res. 16, 4299-4313.
6. Myers, R.M., Tilly, K., and Maniatis, T. (1986) Science 232, 613-618.
7. Cowie, A., and Myers, R.M. (1988) Mol. Cell. Biol. 8, 3122-3128.
8. Treisman, T., Orkin, S.H., and Maniatis, T. (1983) Nature 302, 591-595.
9. Kazazian, H.H., Jr., and Boehm, C.D. (1988) Blood 72, 1107-1116.
10. Mantovani, R., Malgaretti, N., Nicolis, S., Ronchi, A., Giglioni, B., and Ottolenghi, S. (1988) Nucleic Acids Res. 16, 7783-7797.
11. Wall, L., de Boer, E., and Grosveld, F. (1988) Genes Dev. 2, 1089-1100.
12. Dignam, J.D., Lebowitz, R.M., and Roeder, R.G. (1983) Nucleic Acids Res. 11, 1475-1489.
13. Xiao, J-H., Davidson, I., Macchi, M., Rosales, R., Vigneron, M., Staub, A., and Chambon, P. (1987) Genes Dev. 1, 794-807.
14. Singh, H., Sen, R., Baltimore, D., and Sharp, P.A. (1986) Nature 319, 151-153.
15. Hooft van Huijduijnen, R.A.M., Bollekens, J., Dorn, A., Benoist, C., and Mathys, D. (1987) Nucleic Acids Res. 15, 7265-7282.