Characterization of the human ephrin-A4 promoter

Else MUNTHE[1] and Hans-Christian AASHEIM

Department of Immunology, The Norwegian Radium Hospital, Montebello, N-0310 Oslo, Norway

Expression of the ephrin-A4 ligand, a family member of ligands binding the Eph receptor tyrosine kinases, is induced after an antigen–receptor stimulation of lymphocytes. To understand the transcription regulation of the ephrin-A4 gene, its promoter was identified and regulating elements were characterized. The ephrin-A4 promoter contains *cis* elements directing the cell-specific expression. By deletion studies, three specific regions, which were contributing to the transcription activity in lymphoid cells, were localized. In one of these regions, an inverted CCAAT box was identified and shown to bind the transcription activator

nuclear factor-Y (NF-Y). The importance of NF-Y binding for the ephrin-A4 promoter activity is shown by a total abrogation of promoter activity after destruction of its binding site. NF-Y binding and activity are also crucially dependent on the integrity of the surrounding sequence. In addition, electrophoretic mobility-shift assay and serial-mutation analysis of the two remaining regulating regions revealed *cis* regulatory elements contributing to the transcription activity of the ephrin-A4 promoter.

Key words: CCAAT box, lymphoid, nuclear factor-Y.

INTRODUCTION

The ephrins are a family of ligands binding the Eph family of receptor tyrosine kinases (RTKs). They can be divided into two subclasses, those that are bound to the cell membrane through a glycosyl-phosphatidylinositol linkage (ephrin-A1–A5) and those that are membrane-spanning molecules (ephrin-B1–B3) [1]. The Eph kinases are the largest known subfamily of RTKs with 14 distinct members. Members of both ephrins and Eph receptors have been identified in a broad range of animals, including mammals, birds, fish, frogs and worms [2]. In addition, one Eph receptor has recently been identified in *Drosophila* [3]. The Eph receptor–ligand interaction leads to a signal cascade in receptor-carrying cells, initiated by autophosphorylation of the cytoplasmic part of the receptor [4]. Recent studies [5,6] have shown that intracellular signalling also occurs in the ligand-carrying cells, both for the class A and class B ligands. The interaction between ephrins and Eph RTKs has been implicated in neural and epithelial morphogenesis [7], axon guidance and fasciculation [8], control of neural-crest cell migration [9], formation of rhombomere boundaries [10], inhibition of cell–cell adhesion [11], angiogenesis [12] and morphogenesis of capillary beds [13]. In addition, deregulated expression of the Eph receptors or ligands may play a role in malignant transformation of cells [14].

The present study focuses on the transcription regulation of the ephrin-A4 gene in the lymphoid cells. Ephrin-A4 was originally identified in a T-lymphoma cell line and shown to be expressed in human-adult spleen, prostate, ovary, small intestine, colon and in the heart, lung and kidney tissues of human foetus [15]. In addition, we have shown that ephrin-A4 mRNA is also found in lymph node and tonsils and the expression is in particular induced in the lymphocytes stimulated through their antigen receptors [16]. The ephrin-A4 gene *EFNA4* consists of four exons and is clustered together with the genes encoding ephrin-A1 and ephrin-A3, and *EFNA1* and *EFNA3*, on chromosome 1 in man and on chromosome 3 in mouse respectively [17]. Ephrin-A4 has been shown to bind to the receptors EphA2–6 and EphA8 [18–20], and to induce phosphorylation of EphA2, EphA5, EphA6 [18] and EphA8 [20].

There is limited information concerning the regulatory sequences and transcription factors required for the expression of Eph RTK and the ephrins. In mouse brain, the zinc-finger transcription factor Krox-20 is a direct transcription activator of EphA4 RTK [21]; the homeobox proteins HOXA1 and HOXB1 activate the expression of EphA2 RTK [22] and HOXA2 regulates expression of EphA7 RTK [23]. All these transcription factors are involved in the segmentational expression pattern of the Eph RTKs in the rhombomeres. The promoter for the mouse EphA8 gene directs the neural cell-specific expression and the minimal promoter consists of five copies of a stimulating protein-1 (Sp1)-binding site [24]. Recently [25], the EphA3 basal promoter has been defined and shown to be active both in EphA3-positive and in EphA3-negative cell lines. In the present study, it was suggested that DNA methylation might be an important mechanism regulating EphA3 transcription [25].

No molecular characterization of ephrin promoters has been described. Previously, mis-expression of the homeobox-containing genes en-1 and en-2 (engrailed homologues) in chicken brain has been shown to induce strongly the expression of ephrin-A2 and ephrin-A5 [26,27]. Further, one recent study [28] suggested that ephrin-A2 is a target gene for the *otx2*-homeobox transcription factor, an important regulator of brain patterning and morphogenesis.

In the present study, the promoter region of the ephrin-A4 gene (*EFNA4*) was characterized in detail and the promoter activity was examined in transiently transfected haemopoietic cell lines. Deletion and mutation studies defined important regions and sequences involved in regulating ephrin-A4 expression. In particular, one transcription factor, nuclear factor-Y (NF-Y), was shown to be binding to the promoter.

EXPERIMENTAL

Cells

The pre-B cell line Reh (CRL 8286; A.T.C.C., Manassas, VA, U.S.A), the mature-B cell line Bjab (a gift from Dr G. Moldenhauer, University of Heidelberg, Heidelberg,

Germany), the myeloid cell line U973 (DSZM Acc 5) and the lymphoblastoid-cell line JY (a gift from Dr Vaclav Horejsi, Prague, Czech Republic) were grown in RPMI 1640 medium supplemented with 10% foetal bovine serum at $37\,^{\circ}C$ in a humidified atmosphere with 5% CO_2. Blood-isolated B cells were stimulated with anti-μ antibodies ($37.5\,\mu g/ml$, $F[ab']_2$ fragment; Dako, Glostrup, Denmark) or the phorbol ester PMA (5×10^{-8} M) for 24 h as described previously [16].

Characterization of 5′-upstream region of the ephrin-A4 gene

We have described previously the exon–intron organization of the ephrin-A4 gene. The genomic clone included exon 1 and the upstream region of the ephrin-A4 gene [16]. The promoter was sequenced (GenBank® accession no. AJ320547) and analysed by the MatInspector program to identify candidate transcriptional regulatory elements [29].

Reporter-gene constructs of the 5′-upstream ephrin-A4 region

The ephrin-A4 genomic clone was used as a template in PCR reactions to amplify both the 5′-upstream region and a series of deleted variants of the 5′-upstream region in front of a firefly luciferase reporter gene in the pGL3-Basic vector (Promega, Madison, WI, U.S.A.) using the primers listed below. An *Mlu*I-restriction site was included in all forward primers and an *Xho*I-restriction site was included in the reverse primers, for cloning purposes. The PCR products were digested with *Mlu*I and *Xho*I, and inserted in the multiple cloning site of pGL3-Basic firefly luciferase plasmid (Promega). All constructs were verified by DNA sequencing. Plasmid DNA used for cell transfection was isolated using a Maxi Prep kit (Qiagen, Hilden, Germany). The forward primers start at the nucleotide numbered in the primers name and the reverse primers ends at the base in the primers name, related to the nucleotide numbering in Figure 1:

F-664: 5′-CCGCAACCACCTCATTACTTC

F-373: 5′-AATCTTCCAAACTTCTAATAAG

F-366: 5′-TAATAAGGCAGTCTACCCTG

F-295: 5′-TGCTGTTGGAATACGTCC

F-262: 5′-TCGCAGTGCGGTGCAGCCT

F-225: 5′-ACCTGGCGCGCGCTGCGG

F-191: 5′-GTTGGCGCTGGGTGGAGAG

F-174: 5′-GTGGAGAGGAGGAGACG

F-140: 5′-GTTATTCAGGGGCCGGGGTCAG

F-111: 5′-ACCCCACTCCCCTTTCC

F-98: 5′-CCTTTCCGCAACTTCCCT

R-13: 5′-CCGGTTTGGTCTGGCCTGGC

R-131: 5′-CCCTGAATAACTTTCCGCC

R-160: 5′-CGTCTCCTCCTCTCCAC

R-265: 5′-GGGCGGGAAGGAAGGACGTA

Site-directed mutagenesis

Mutations were introduced in the ephrin-A4 promoter constructs using the QuikChange Site-Directed Mutagenesis kit from Stratagene (La Jolla, CA, U.S.A.) according to the manufacturer's instructions, to define the elements important for the observed electrophoretic mobility shifts and reporter-gene activation. Two elements, the CCAAT box and the putative myb site in the F664-13 construct, were mutated. Primers used for the PCR to mutate the CCAAT box were: F: 5′-GGAGACGCGC-TCCCAT**CC**GCGGAAAGTTATTC and R: 5′-ATAACTTT-CCGC**GG**ATGGGAGCGCGTCTCC.

Primers used in PCR to mutate the putative myb site were: F: 5′-GCTCCCATTGGCGGAAAG**CCC**TTCAGGGGCCG-GGGTCAG and R: 5′-CTGACCCCGGCCCCTGAA**GGG**CT-TTCCGCCAATGGGAGC.

The mutated bases are underlined and are shown in boldface. The mutated fragments were re-cloned into the pGL3 vector and the introduction of mutations were verified by DNA sequencing.

Linker scanning analysis

To define the binding sites in regions I and II, mutations were introduced using the following upstream PCR primers, in combination with the R13 primer described above. The mutated bases, as indicated in Figures 6 and 7, are here shown in boldface:

F321: 5′-CGAAAGAGATGACCTCCCTAAAAATACTGC

F321mut1: 5′-CGA**CCTC**GATGACCTCCCTAAAAATACT-GC

F321mut2: 5′-CGAAAGAG**CCC**ACCTCCCTAAAAATACT-GC

F321mut3: 5′-CGAAAGAGATGA**AAGA**CCTAAAAATAC-TGC

F321mut4: 5′-CGAAAGAGATGACCTC**AAGC**AAAATACT-GC

F321mut5: 5′-CGAAAGAGATGACCTCCCTA**CCCC**TACT-GC

F225: 5′-ACCTGGCGCGCGCTGCGGTGCACAGAGGGT-TAACTGGA

F225mut1: 5′-ACCTGGCGCGCGC**CCAA**GTGCACAGAG-GGTTAACTGGA

F225mut2: 5′-ACCTGGCGCGCGCTGCG**CCCG**ACAGAG-GGTTAACTGGA

F225mut3: 5′-ACCTGGCGCGCGCTGCGGTGC**TTCC**AG-GGTTAACTGGA

F225mut4: 5′-ACCTGGCGCGCGCTGCGGTGCACA**CGC**-GGTTAACTGGA

F225mut5: 5′-ACCTGGCGCGCGCTGCGGTGCACAGAG-**ACG**TAACTGGA

F225mut6: 5′-ACCTGGCGCGCGCTGCGGTGCACAGAG-GGT**GCC**CTGGA

An *Mlu*I-restriction site was included in all the forward primers for cloning purposes. The PCR products were digested with *Mlu*I and *Xho*I, and inserted in the multiple cloning site of pGL3-Basic firefly luciferase plasmid (Promega). All constructs were verified by DNA sequencing.

Transient transfection and luciferase assay

Transfection of cells was performed by electroporation. Exponentially growing cells (0.5–1 million/ml) were harvested and

resuspended in RPMI 1640 and kept on ice. Equal amounts (5 μg) of the different ephrin-A4 promoter plasmids (full-length promoter, deletion construct and mutated construct), the positive transfection control pGL3-Control plasmid (SV40 promoter; Promega) and the promoterless negative transfection control pGL3-Basic (Promega), along with 3 μg of the internal transfection control pRL-CMV (cytomegalovirus) plasmid (Promega), were added to 400 μl of cells (1–2 millions/ml). Electroporation was conducted at 200 V for 75, 10, 50 and 20 min for Reh, Bjab, U937 and JY cells respectively, using an ECM 830 Electro Square Porator (BTX Inc., San Diego, CA, U.S.A.). Immediately after electroporation, 800 μl of conditioned medium, supplemented with 5% foetal bovine serum was added and the cells were incubated at 37 °C. Luciferase activities in transfected cells were measured twice, 8–24 h after electroporation, using the FireLite kit (Packard, Groningen, The Netherlands) and a Top Count liquid-scintillation counter (Packard, Meriden, CT, U.S.A.).

Electrophoretic mobility-shift assay (EMSA)

Essential nuclear extracts were prepared as described by Schreiber et al. [30]. In brief, $(10–50) \times 10^6$ cells were washed with PBS, resuspended in 800 μl ice-cold buffer A [10 mM Hepes (pH 7.9), 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 1 mM 1,1,1-trichloro-2,2-bis-(p-chlorophenyl)ethane (DDT), 0.5 mM PMSF and 0.5% of phosphatase inhibitor cocktail II (Sigma)], then allowed to swell on ice for 15 min before lysis by addition of 50 μl of 10% Nonidet P40 and vigorous vortex-mixing. Cell nuclei were pelleted in a microfuge at 6000 rev./min (3300 g) for 2 min, resuspended in 50 μl of ice-cold buffer C [20 mM Hepes (pH 7.9), 0.4 M NaCl, 1 mM EDTA, 1 mM EGTA, 1 mM DDT, 1 mM PMSF and 0.5% of phosphatase inhibitor cocktail II] and incubated on ice for 30 min with occasional flicking. The nuclei were pelleted by centrifugation and aliquots of the supernatants were stored at −20 °C. Protein concentrations were estimated by Ponceau Red (Sigma) staining of dot-blots on nitrocellulose membranes (Schleicer and Schuell GmbH, Dassel, Germany) by comparison with proteins of known concentration. Blunt double-stranded DNA fragments were generated by PCR using the three primer pairs F373 and R265, F225 and R160, and F191 and R131 as described above, for the first, second and third binding sites respectively, and end-labelled with [γ-^{32}P]ATP using T4 polynucleotide kinase (Promega).

Labelled DNA was incubated with 2 μg of the nuclear extract at 20 °C for 25 min with 0.3 μg of poly(dI/dC) (Pharmacia, Uppsala, Sweden) in 10 μl of a buffer containing 10 mM Hepes (pH 7.9), 100 mM KCl, 0.2 mM EDTA, 1 mM DDT and 2% (v/v) glycerol. Unlabelled competitor DNA or 2 μl of antibody raised against the A subunit of the transcription factor NF-Y (Rockland Immunochemicals, Gilbertville, PA, U.S.A.) or an irrelevant antibody raised against ephrin-B1 (Santa Cruz Biotechnologies) were added 20 min before the addition of labelled DNA. Two complementary oligonucleotides (sense 5′-CGTCTCCACCAATGGGAGGGCTGGGC-3′), including the CCAAT box from the CD10 promoter, were annealed to investigate if the retardation could be competed out using an unrelated CCAAT box [31]. Two complementary oligonucleotides (sense 5′-GATCGCTCTAAAAATAACCCTGTCG-3′), including an MEF2-binding site from the muscle creatine kinase (MCK) promoter were annealed to investigate if the retardation could be competed out using an unrelated CCAAT box [32].

To examine the CCAAT-box surrounding sequences, complementary mutated oligonucleotides were annealed and used as competitors. The following sense sequences were applied, with the mutated bases underlined and indicated in boldface:

competitor 1: 5′-GCGCTCCCATTGGCGGAAAGTTATTCA-GGGGCC

competitor 2: 5′-GCGCTCCCATTGGCGGAAAG**CCC**TTCA-GGGGCC

competitor 3: 5′-GCGCTCCCAT**CC**GCGGAAAGTTATTC-AGGGGCC

competitor 4: 5′-GCGCTCCCATTGGCGGAAA**T**TTATTC-AGGGGCC

competitor 5: 5′-GCGCTCCCATTGGCGGAAAGTTA**CCC**A-GGGGCC

After incubation of nuclear extracts with labelled DNA, the samples were separated on 4% polyacrylamide gels (acrylamide/bis-acrylamide, 29:1) in 0.5 × Tris/borate/EDTA or 1 × GR buffer [50 mM Tris/HCl, 400 mM glycine and 2 mM EDTA (pH 7.5)] at 4 °C and 75 V. Gels were exposed and analysed with a PhosphorImager (Molecular Dynamics, Pharmacia).

UV cross-linking

An up-scaled EMSA was performed and the retarded band was isolated. The gel slice was divided and half of it was UV irradiated for 20 min on ice. The gel slices were then solubilized in 20 μl of 3 × SDS buffer, boiled for 5 min and separated by SDS/PAGE (10% gel). The gel was exposed and analysed with a PhosphorImager.

RESULTS

Characterization of the 5′-region of the ephrin-A4 gene

We have previously presented the genomic organization of the human ephrin-A4 gene consisting of four exons and spanning more than 7 kb [16]. The genomic clone originally isolated from a normal male peripheral leucocyte genomic DNA library, covered the entire coding region of the ephrin-A4 gene in addition to approx. 600 bp of its promoter. The sequence of the 5′-untranslated region (5′-UTR) and promoter (GenBank® accession no. AJ320542) with putative regulatory elements, as defined by the MatInspector program [29], using the TransFac database [33], is presented in Figure 1.

No consensus TATA box is present in the ephrin-A4 promoter. An AT-enriched stretch is localized from nucleotides −144 to −136, harbouring a putative myb site. An inverted CCAAT box is localized 150 nucleotides upstream of the translation-initiation site. Several transcription factors are known to bind the CCAAT boxes. These include NF-Y, CCAAT-enhancer-binding protein, nuclear factor-1 and CCAAT-displacement protein [34]. An AT-enriched stretch, matching a consensus myocyte enhancer factor-2 (MEF2)-binding site is identified at nucleotides −305 to −297 [35]. Several consensus GATA elements are identified and also one classical GC box is present (nucleotides −271 to −263), known to be a potential binding site for the SP family of transcription factors. In addition, a GC-rich island spans parts of the promoter as well as the first exon.

To compare the promoter sequences with different cell sources, we amplified the promoter region by PCR on genomic DNA from peripheral blood leucocytes and the Daudi cell line. Comparison of these sequences, as well as comparison with a partial genomic sequence from GenBank® (accession no. Z57588), did

```
-679  GGTGCTTTGTACAGGCCGCAACCACCTCATTACTTCGTCTTAGGGACTGG

-629  GGCCGCGTGGGCCCCCAGCCCGGAACGAAGGTGTGGAGCGGCAAGGGACA

-579  GACGCCAATCCTAAAGTGAGCATCTAGCGCGCCTCCTAAGGCTCTTTAGG

-529  GAAGGTGGTCCCAGAGCTGTGTTGTCCCTTCCGCTTGCACTGTCCCTAGA

-479  TGTGCAAAGAAAACGGGGCAGTGCATGAAGGTGGTTGGACAGGCTTCAGG

-429  ATCCTCGCCCGCGCCTCACTTTCCCCTATCTGGGCAAAGGTTATGTACCC
                    ‾‾‾‾‾‾
                     GATA

-379  TTATTTAAAATCTTCCAAACTTCTAATAAGGCAGTCTACCCTGCACTAAA
            ‾‾‾‾‾‾
             GATA

-329  GCAGACACGAAAGAGATGACCTCCCTAAAAATACTGCTGTTGGAATACGT
                           ‾‾‾‾‾‾‾‾      ‾‾‾‾‾‾
                             MEF-2          C-MYB

-279  CCTTCCTTCCCGCCCCCCTCGCAGTGCGGTGCAGCCTCAGTGGAAGCTTTG
              ‾‾‾‾‾‾‾‾‾‾
               GC-box

-229  GCGAACCTGGCGCGCGCTGCGGTGCACAGAGGGTTAACTGGAGTTGGCGC
                   ‾‾‾‾‾‾
                    AP4

-179  TGGGTGGAGAGGAGGAGACGCGCTCCCATTGGCGGAAAGTTATTCAGGGG
                                  ‾‾‾‾‾     ‾‾‾‾‾‾
                                  NF-Y        C-MYB

-129  CCGGGGTCAGTGAATCTCCGTACCCCACTCCCCTTTCCGCAACTTCCCTC

-79   TTCACTTTGTACCTTTCTCTCCTCGACTGTGAAGCGGGCCGGGACCTGCC

                                               +1
-29   AGGCCAGACCAAACCGGACCTCGGGGGCGATG
```
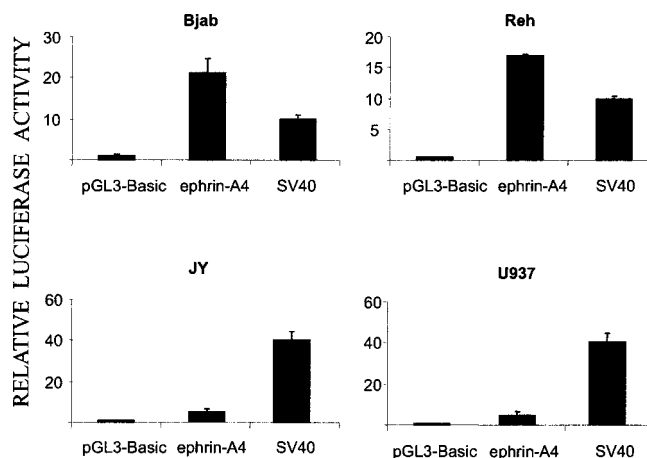
Figure 1   The ephrin-A4 promoter and 5′-UTR

The sequence of the human ephrin-A4 upstream region is presented (GenBank® accession no. AJ320547). The adenine in the translation-initiation codon (ATG) is denoted as +1. The sequence was analysed with the MatInspector program [29] and putative binding sites for transcription factors are indicated (underlined).

not identify any polymorphism in the regulatory regions of the ephrin-A4 promoter (results not shown).

Ephrin-A4 upstream region contains elements necessary for cell-specific regulation

The 5′-upstream region of the human ephrin-A4 gene was analysed for its transcription activity when fused to a promoterless firefly luciferase gene. This region covered the sequence from nucleotides −664 to −13 relative to the start codon and included the 5′-UTR (Figure 1). This promoter construct, hereafter denoted as the full-length promoter, was tested for its ability to induce expression of the reporter gene in transiently transfected haemopoietic cell lines. The cells were co-transfected with an internal transfection control, a vector containing the *Renilla* luciferase gene under control of the CMV promoter. Two B cell lines, Reh and Bjab, expressing high levels of ephrin-A4 mRNA, two cell lines, JY and U937, expressing low levels of ephrin-A4 mRNA [16] were transfected with the full-length ephrin-A4 promoter-luciferase construct. The results are presented in Figure 2. The ephrin-A4 promoter induced 30- and 18-fold up-regulation of luciferase activity in Reh and Bjab cells respectively, when compared with the promoterless vector-control (pGL3-Basic) activity, and with even higher activity than an SV40-positive control vector. The ephrin-A4 promoter induced only a 6- and 5-fold increase in luciferase activity respectively in JY and U937 cells, when compared with the promoterless vector-control activity. These results suggest that the promoter contains

Figure 2   The ephrin-A4 promoter contains *cis* elements necessary for cell-specific activity

The ephrin-A4 promoter in front of a firefly luciferase reporter gene (ephrin-A4) was transiently co-transfected in Reh, Bjab, JY and U937 cells with a control plasmid containing the *Renilla* luciferase gene under control of the CMV promoter. A positive control vector, SV40 promoter in front of firefly luciferase reporter gene (SV40), was co-transfected with the *Renilla* luciferase gene in the same cell lines. A negative control vector (pGL3-Basic) was co-transfected with the *Renilla* luciferase gene in the same cell lines. Relative luciferase activities are expressed as firefly luciferase activity normalized to the *Renilla* luciferase activity. Shown here are means of one representative experiment performed in triplicate. Bars show the standard deviation. At least three different experiments for each cell line were performed.

A

B
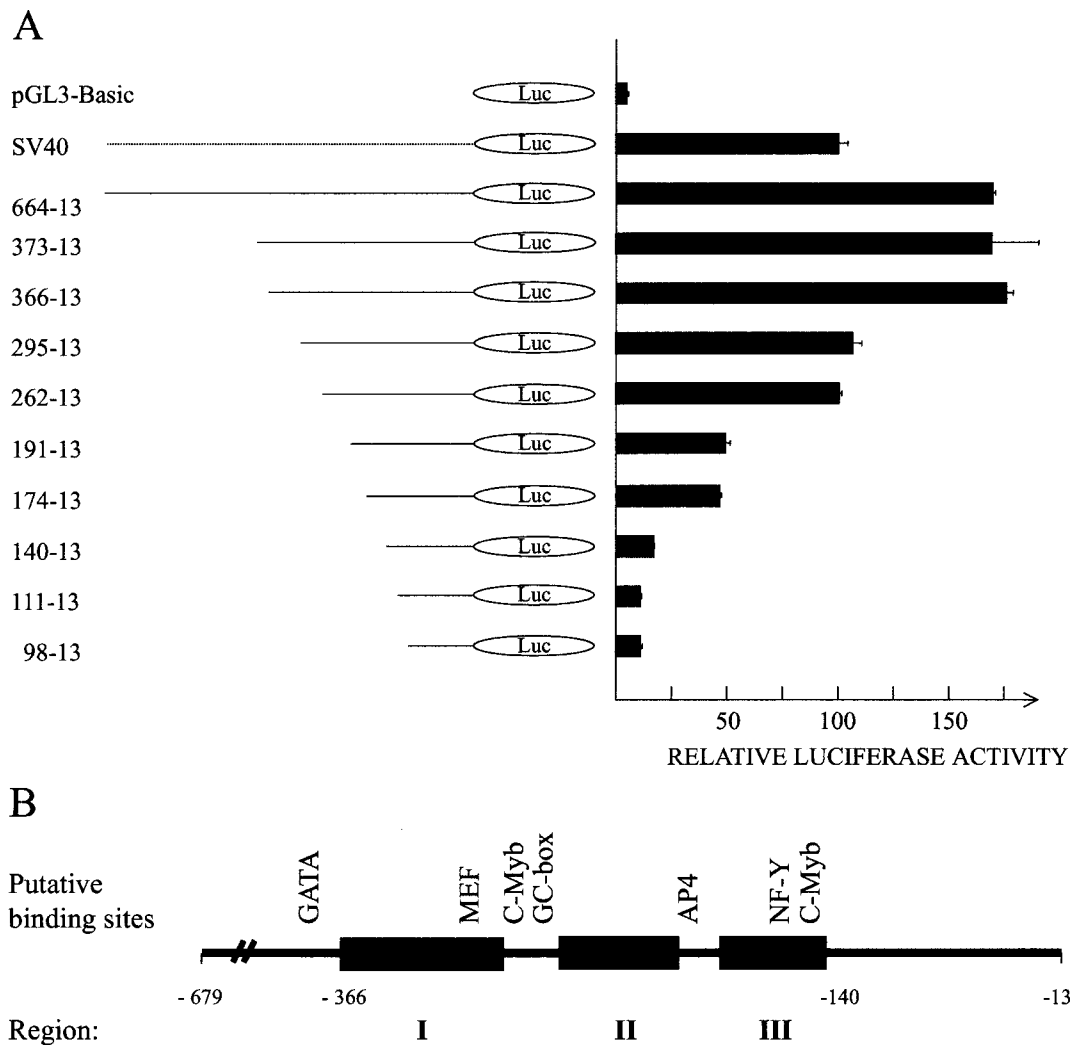
**Figure 3   Identification of *cis* elements important for the ephrin-A4 expression**

(**A**) Progressive shorter ephrin-A4 promoter constructs in front of a firefly luciferase reporter gene were transiently co-transfected in Reh cells with a control plasmid containing the *Renilla* luciferase gene under control of the CMV promoter. A positive control vector, SV40 promoter, in front of firefly luciferase reporter gene (SV40) was co-transfected with the *Renilla* luciferase gene. A negative control vector, pGL3-Basic, was co-transfected with the *Renilla* luciferase gene in the same cell lines. Relative luciferase activities are expressed as firefly luciferase activity normalized to the *Renilla* luciferase activity. The construct numbers shown on the left-hand side of the Figure reflect the nucleotide position in the promoter. Shown here are means of one representative experiment performed in triplicate. Bars show the standard deviation. At least three different experiments for each construct were performed. (**B**) The experiments defined three distinct regions affecting transcription of the ephrin-A4 gene. An overview of the localization of these three regions in the ephrin-A4 promoter is shown and putative transcription factor-binding sites are indicated.

the *cis* elements necessary for cell-specific transcription of the ephrin-A4 gene.

Identification of regulatory elements in the ephrin-A4 promoter

A series of deletion constructs was made to identify the specific regions involved in regulating the transcription activity of the ephrin-A4 gene. Progressive deletions from nucleotides −664 to −13 were fused to the luciferase-reporter gene (see Figure 3A). Transient transfection of these constructs in Reh cells showed that the sequence between nucleotides −366 and −140 was required for maximal promoter activity (Figure 3A). Deletion from nucleotides −366 to −295 decreased the luciferase activity from 30- to 22-fold over the background, represented by the activity directed from the promoterless pGL3-Basic vector. This region, denoted as region I (nucleotides −366 to −295), includes the consensus-binding sites for the transcription factors MEF2 and c-*myb* as

revealed by the MatInspector program [29]. Further deletion to nucleotide −262 did not decrease the luciferase activity, but a decrease in luciferase activity to 10-fold over the background was obtained by deletion to nucleotide −191. The data-program analysis did not reveal any consensus-binding sites in this second region (region II, nucleotides −261 to −191). Further deletion to nucleotide −174 did not decrease the luciferase activity, but deletion to nucleotide −140 decreased the promoter activity to only 3-fold over the background level. This third region (region III, nucleotides −174 to −140) includes an inverted CCAAT box and an AT-enriched stretch with a putative c-*myb*-binding site. Interestingly, no significant difference in expression was observed between the two constructs covering the nucleotides from −295 to −13 and from −262 to −13 (Figure 3A), indicating that a GC box (localized between nucleotides −271 and −263) is not involved in regulating the basal expression of the ephrin-A4 gene. A schematic overview of the localization of

© 2002 Biochemical Society
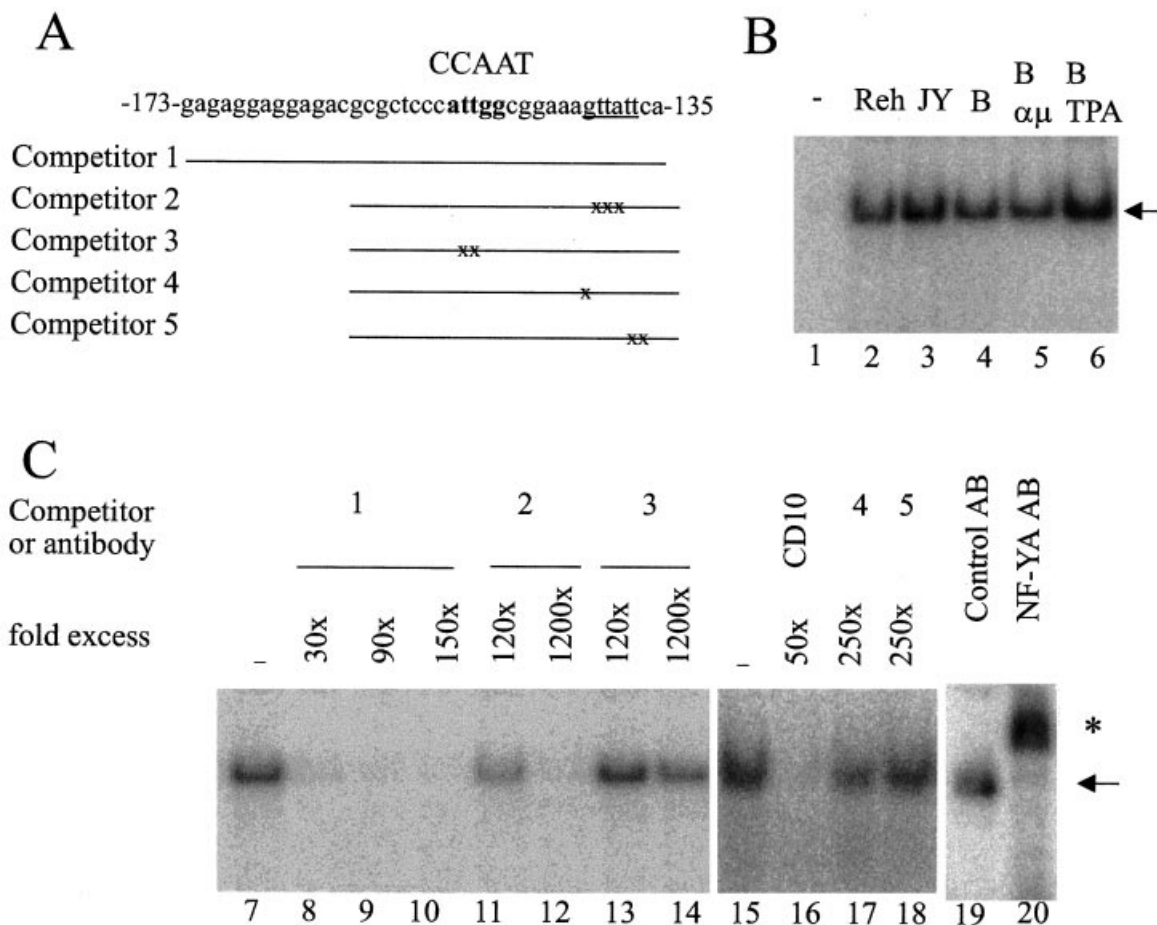
Figure 4   The CCAAT box in region III binds the transcription factor NF-Y

(**A**) A schedule of the region III DNA fragment used as probe, including the inverted CCAAT box and the putative myb site and a schematic overview of the ephrin-A4 competitors used in the EMSA. The inverted CCAAT box is shown in boldface and a part of the myb site and two T nucleotides that might be involved in the binding of NF-Y to region III are underlined. The x denotes the mutated nucleotides described under the Experimental section. (**B**) EMSA of nuclear extracts. 1. End-labelled probe (region III fragment) was incubated with 5 μg nuclear lysates from: lane 1, no lysate; lane 2, Reh; lane 3, JY; lane 4, B cells; lane 5, anti-IgM-stimulated B cells; lane 6, PMA-stimulated B cells. (**C**) The CCAAT box and the surrounding sequence is essential for the binding of the transcription factor NF-Y in region III. End-labelled wild-type fragment (lanes 7–20) was incubated with the following excess of unlabelled competitor or antibody for 20 min before the addition of 5 μg of nuclear Reh lysate; lane 8, 30-fold excess competitor 1; lane 9, 90-fold excess competitor 1; lane 10, 150-fold excess competitor 1; lane 11, 120-fold excess competitor 2; lane 12, 1200-fold excess competitor 2; lane 13, 120-fold excess competitor 3; lane 14, 1200-fold excess competitor 3; lane 16, 50-fold excess CD10 promoter CCAAT box; lane 17, 250-fold excess competitor 4; lane 18, 250-fold excess competitor 5; lane 19, +anti-ephrin-B1; lane 20, +anti-NF-Y antibody. No competitor was added in lanes 7 and 15. Arrows in (**B**) and (**C**) indicate the specific retarded band and * denotes the super-shifted complex. Fold excess of the competitors are indicated above the Figure and the numbers denote the competitors used in the EMSA reaction.

the three defined regions affecting ephrin-A4 transcription is presented in Figure 3(B).

Transcription activator NF-Y binds to the inverted CCAAT box in the ephrin-A4 promoter

EMSA was performed to identify the proteins binding to the three important putative regions (I, II and III) identified by the serial-deletion experiments described above (Figure 3). EMSA was performed with nuclear extracts isolated from the Reh and JY cell lines, and with those isolated from resting blood B lymphocytes and B lymphocytes activated via their antigen receptor or with the phorbol ester PMA (Figure 4B). A single, strong shift of a ubiquitous nature was observed using a ^{32}P-labelled probe covering the 3′-part of region III (nucleotides −173 to −135), including the inverted CCAAT box and the surrounding se-

quence. The specificity of this retardation (Figure 4C, lane 7) was proven by the elimination of the shift using increasing excess of the identical unlabelled fragment in the reaction (Figure 4C, lanes 8–10).

We suspected the CCAAT box to be responsible for this shift by binding a specific transcription factor. Therefore a fragment with mutations in the inverted CCAAT box (from ATTGG to ATCCG, competitor 3 in Figure 4A) was generated and used as competitor in the EMSA. The result from this experiment shows that a 120-fold excess of the mutated fragment did not affect the retardation and even a 1200-fold excess of mutated fragment did not eliminate the retardation obtained with the wild-type fragment (Figure 4C, lanes 13 and 14). A shorter wild-type fragment corresponding to the competitors was able to eliminate the shift (results not shown). To confirm further the importance of the CCAAT box, annealed oligonucleotides covering the CCAAT
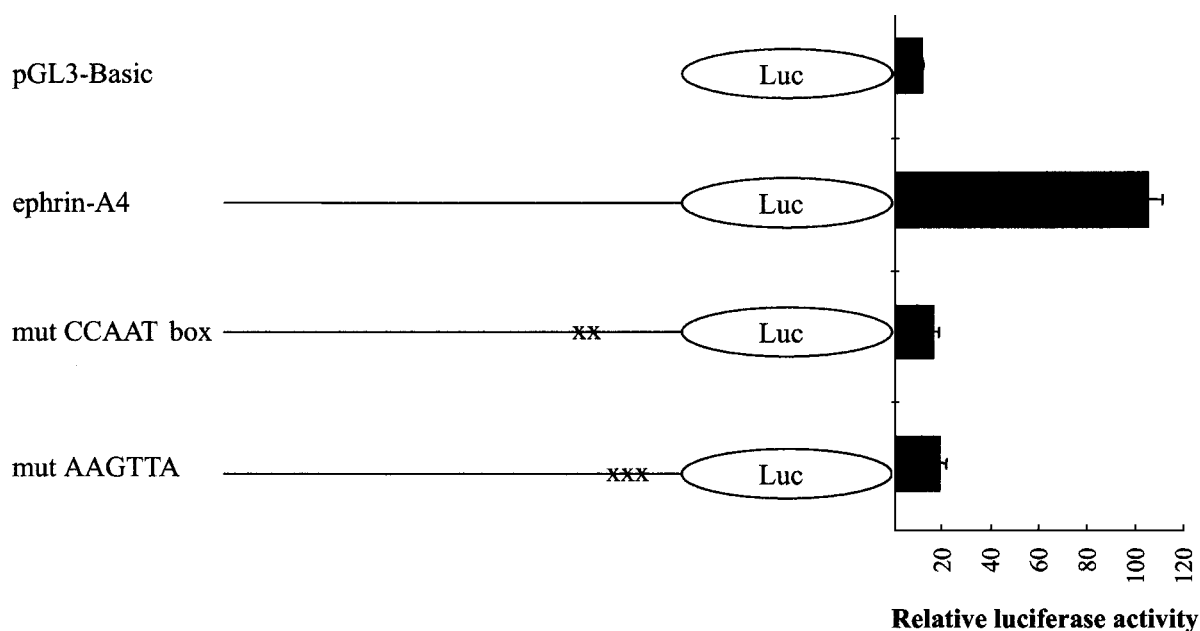
Figure 5   Both CCAAT box and putative myb site are necessary for the promoter activity

Reh cells were transiently transfected using a negative control vector pGL3-Basic, the wild-type ephrin-A4 promoter (ephrin-A4) or ephrin-A4 promoters with mutated CCAAT box (mut CCAAT box; from ATTGG to ATCCG) or mutated putative myb site (mut AAGTTA; from AAGTTA to AAGCCC). The specific mutations are denoted by 'x' and are described in the Experimental section. A control plasmid was co-transfected in all cases, containing the *Renilla* luciferase gene under the control of CMV promoter. Relative luciferase activities are expressed as firefly luciferase activity normalized to the *Renilla* luciferase activity. Shown here are means of a representative experiment performed in triplicate. Bars show the standard deviation. Three different experiments were performed. Mutations in the CCAAT box and in the putative myb site (AAGTTA), both abrogate the promoter activity.

box from a different promoter, the CD10 promoter [31], was added to the EMSA reaction. A 50-fold excess of the CD10 promoter derived from the CCAAT box successfully eliminated the retardation complex obtained with the wild-type fragment (Figure 4C, lane 16). The conclusion to be drawn from these experiments is that the inverted CCAAT box is necessary for the binding of a transcription factor to region III in the ephrin-A4 promoter.

Next, we examined whether the ubiquitously expressed factor NF-Y is able to bind the ephrin-A4 promoter. NF-Y was a strong candidate since the CCAAT-box sequence in the ephrin-A4 promoter matched the consensus-binding site for the heterotrimeric NF-Y factor [34,36], in addition to the observed ubiquitous nature of the region III DNA fragment retardation (Figure 4B). Addition of an antibody specific to the A subunit of NF-Y in the EMSA reaction, resulted in a supershift of the retarded complex (Figure 4C, lane 20), confirming that NF-Y has the ability to bind the CCAAT box in the ephrin-A4 promoter.

Binding of NF-Y to the inverted CCAAT box is dependent on the integrity of the surrounding sequence

A putative consensus myb-binding site (AAGTTA) located downstream from a CCAAT box in the 5′-UTR of the gene encoding the cell-surface protein CD34 has previously been shown to be important for the binding of the transcription factor NF-Y [37]. In the present study, it was demonstrated that the integrity of the putative myb site is involved in stabilizing the binding of NF-Y to its site and also that it is important for the

transcription activity of the CD34 promoter in transfected cells. An identical putative myb-binding site (AAGTTA) is also present downstream of the CCAAT box in the ephrin-A4 promoter (nucleotides −143 to −138; see Figures 1 and 4). We therefore evaluate whether this site had any impact on the binding of the NF-Y to the ephrin-A4 promoter. An unlabelled fragment carrying mutations in the putative myb site (from AAGTTA to AAGCCC; competitor 2 in Figure 4A) was added in the EMSA reaction together with the end-labelled wild-type region III fragment. A 120-fold excess of this unlabelled mutated fragment (competitor 2) slightly decreases the retardation (cf. lane 7 with lane 11 in Figure 4C), whereas a 1200-fold excess completely eliminates the retardation (cf. lane 7 with lane 12 in Figure 4C). In addition, a mutation from G to T, in the putative myb site (AAGTTA to AATTTA; competitor 4, Figure 4A) was introduced, both to destroy the putative myb site and to improve the AT-enriched stretch located 3′ for the CCAAT box. No decrease in the retardation of the wild-type fragment was observed with 250-fold excess of competitor 4 (Figure 4C, lane 17) showing that G in the putative myb site is involved in the NF-Y binding. This result also indicates that the AT-enriched stretch by itself is not crucial for the binding of NF-Y. Further, a 250-fold excess of a fragment with a TT to CC mutation downstream of the myb site (competitor 5, Figure 4A), which also shortens the AT-enriched stretch, did not reduce the retardation of the wild-type fragment (Figure 4C, lane 18). This indicates that GTTATT (part of the putative myb site and two T nucleotides) might be involved in NF-Y binding to region III.

To conclude, a relationship exists between the inverted CCAAT box and the surrounding sequence for stable binding of the NF-Y complex to region III in the ephrin-A4 promoter.
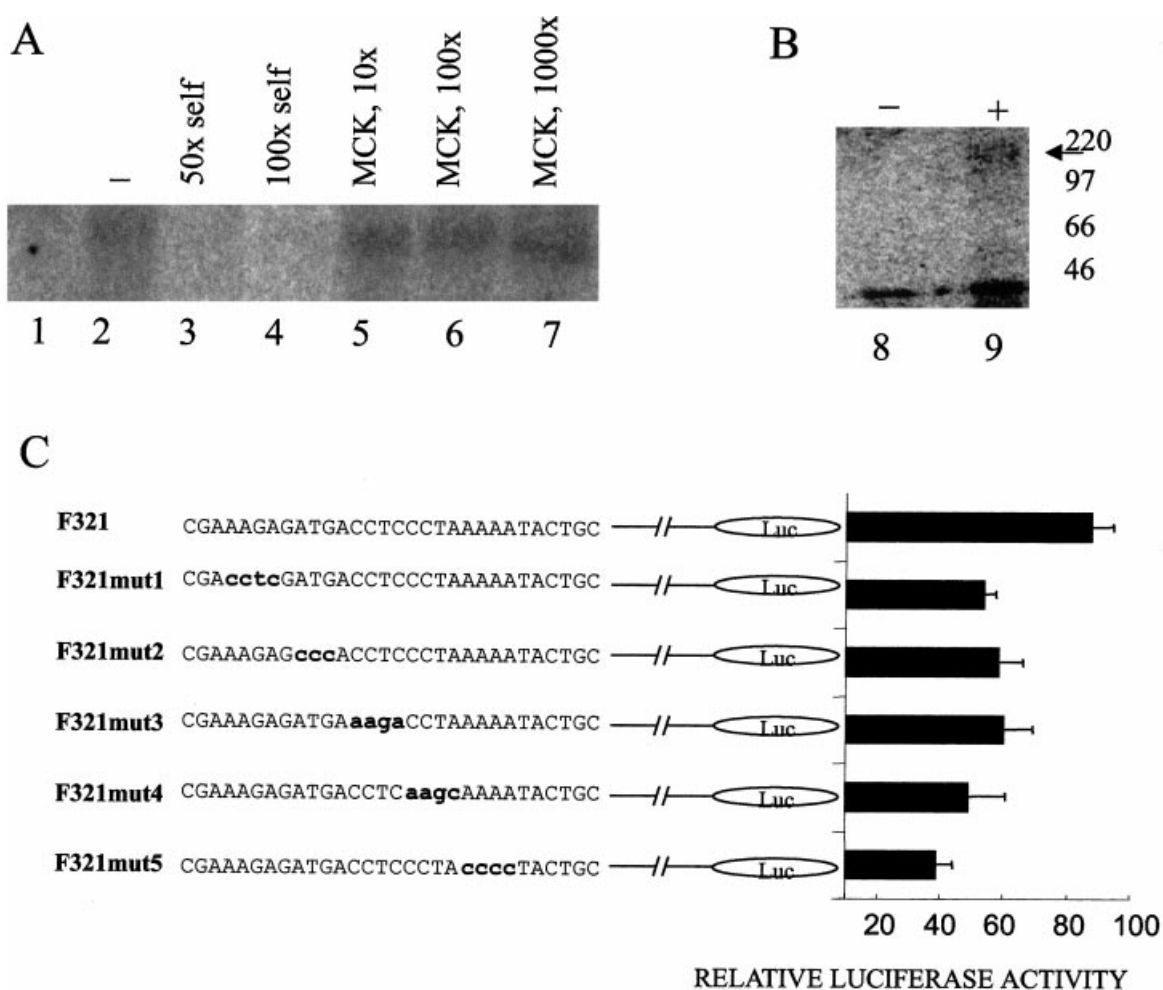
Figure 6   Identification of a nuclear factor binding to region I

(**A**) The end-labelled probe covering the region I was incubated with Reh nuclear lysate and the indicated competitors. Lane 1, no lysate; lane 2, Reh lysate; lanes 3 and 4, 50 and 100 times excess unlabelled probe respectively; lanes 5–7, 10, 100 and 1000 times excess annealed MCK primers. (**B**) The retarded complex from an up-scaled EMSA was isolated and divided into two out of which one part was UV cross-linked. Lane 8, no UV cross-linking; lane 9, UV cross-linking for 20 min. Arrow indicates the UV cross-linked DNA and protein. Indicated to the left-hand side of the Figure are the molecular masses in kDa. The molecular mass of the labelled probe is approx. 40 kDa. (**C**) A linker scanning analysis revealed the important nucleotides affecting transcription in region I. The sequences are shown on the left-hand side of the Figure. Wild-type ephrin-A4 region II is denoted as F321 (from nucleotides −321 to −13) and mutations are denoted as F321 mut1–5. The small boldface letters in the sequences indicate the mutations. The ephrin-A4 promoter constructs in front of a firefly luciferase reporter gene were transiently co-transfected in Reh cells with a control plasmid containing the *Renilla* luciferase gene under the control of the CMV promoter. Relative luciferase activities (to the right-hand side in the Figure) are expressed as the firefly luciferase activity normalized to the *Renilla* luciferase activity. Mutations between nucleotides −318 and at least −296 (F321 mut1–5) inhibit the ephrin-A4 promoter activity.

Both the inverted CCAAT box and the putative myb-binding site is essential for the ephrin-A4 promoter activity in transfected cells

Next, we investigate whether the inverted CCAAT box is involved in regulating the ephrin-A4 promoter-induced transcription in transfected cells. Mutations were introduced in the CCAAT box, from ATTGG to ATCCG, in the ephrin-A4 promoter–luciferase construct. A dramatic reduction in the relative luciferase activity was observed after transient transfection of Reh cells with the mutated construct when compared with the wild-type full-length promoter (Figure 5). A decrease from 30- to 3-fold, over the promoterless-luciferase construct to the same level, was observed when the CCAAT box was removed (Figure 3, 111–113 construct). Further, Reh cells were transiently transfected with an ephrin-A4 promoter construct carrying mutations in the
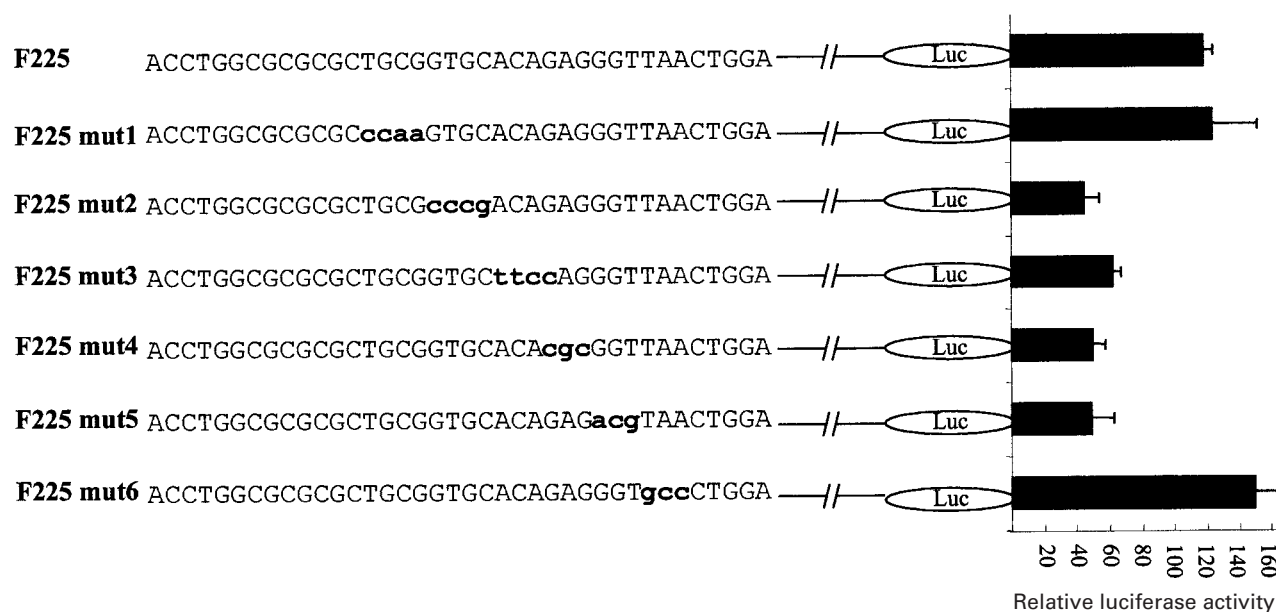
putative myb site (from AAGTTA to AGGCCC) to investigate whether this site is also involved in regulating ephrin-A4 promoter-induced transcription. A dramatic decrease in the luciferase activity was also observed for this construct when compared with the wild-type full-length promoter, which was comparable with the decrease in activity observed with the mutated CCAAT box (Figure 5).

To conclude, the CCAAT box and the integrity of the surrounding sequence are essential for the transcription activity of the ephrin-A4 gene.

Identification of the regulatory sequences in regions I and II

Two additional regions (regions I and II) strongly contribute to the transcription activity of the ephrin-A4 promoter shown by

F225   ACCTGGCGCGCGCTGCGGTGCACAGAGGGTTAACTGGA —//— Luc

F225 mut1 ACCTGGCGCGCGC**ccaa**GTGCACAGAGGGTTAACTGGA —//— Luc

F225 mut2 ACCTGGCGCGCGCTGCG**cccg**ACAGAGGGTTAACTGGA —//— Luc

F225 mut3 ACCTGGCGCGCGCTGCGGTGC**ttcc**AGGGTTAACTGGA —//— Luc

F225 mut4 ACCTGGCGCGCGCTGCGGTGCACA**cgc**GGTTAACTGGA —//— Luc

F225 mut5 ACCTGGCGCGCGCTGCGGTGCACAGAG**acg**TAACTGGA —//— Luc

F225 mut6 ACCTGGCGCGCGCTGCGGTGCACAGAGGGT**gcc**CTGGA —//— Luc

Relative luciferase activity

Figure 7   Identification of the important *cis* element in region II

A linker scanning analysis reveals the important nucleotides affecting transcription in region II. The sequences are shown on the left-hand side of the Figure. Wild-type ephrin-A4 region II is denoted as F225 (from nucleotides −225 to −13) and mutations are denoted as F225 mut1–6. The small boldface letters in the sequences indicate the mutations. The ephrin-A4 promoter constructs in front of a firefly luciferase reporter gene was transiently co-transfected in Reh cells with a control plasmid containing the *Renilla* luciferase gene under the control of the CMV promoter. Relative luciferase activities (to the right-hand side of the Figure) are expressed as the firefly luciferase activity normalized to the *Renilla* luciferase activity. Mutations between nucleotides −208 and −196 (F225 mut2–5) inhibit the ephrin-A4 promoter activity.

deletion analysis of the ephrin-A4 promoter (Figure 3). EMSA was performed with a fragment covering region I (nucleotides −366 to −265; Figure 6A) to define the proteins binding this region in nuclear extracts from Reh cells. In the present study, retardation was observed (Figure 6B). The retardation obtained with the labelled region I fragment was competed out with a 50-fold excess of the same unlabelled fragment, indicating the specificity of this retardation. Next, the specific EMSA band was isolated and UV-irradiated to cross-link the end-labelled DNA to protein to determine the size of the binding factor. A band of approx. 200 kDa was identified, which appeared only after UV cross-linking (Figure 6C). The weight of the DNA probe was approx. 40 kDa, indicating that the molecular mass of the binding factor(s) should be more than 150 kDa. As mentioned above, region I contains putative binding sites for MEF2 (55 kDa). The observed molecular mass indicates that if MEF2 binds to region I, it will have to appear in stable complexes. To examine further the putative MEF2-binding site, we tested if an oligonucleotide covering the MEF2-binding site from the MCK enhancer, shown to eliminate a specific MEF2 shift in lymphocytes [32], was able to eliminate the shift in region I. No effect on the retardation was observed using this oligonucleotide (Figure 6A), which indicated that MEF2 might not bind to region I. To narrow down the interesting area further, a deletion analysis was performed (results not shown). The results showed that the nucleotides between −321 and −265 were involved in the activities of region I. Further, a series of sequential mutations, a so-called linker scanning analysis, was performed to reveal the nucleotides, which contribute to the promoter activity in region I. The nucleotides between −321 and −13 were included in the promoter constructs, and an overview of the mutations is shown in Figure 6(C). The results from this analysis showed that all the introduced

mutations inhibited the reporter-gene activity, suggesting that both the pyrimidine stretch and nucleotides upstream for this stretch are involved in the transcription activity from region I.

A deletion analysis was performed to narrow region II (results not shown). This analysis revealed that nucleotides distal from −225 did not contribute to the activity. No shift was detected with a probe that covered region II (nucleotides −225 to −160; results not shown). A linker scanning analysis was then performed to reveal nucleotides that contribute to the promoter activity in region II. The nucleotides between −225 and −13 were included in the promoter constructs, and an overview of the introduced mutations is shown in Figure 7. The results from this analysis showed that the introduced mutations numbered 2–5, abrogated the contribution of region II for the ephrin-A4 promoter activity, whereas mutations numbered 1 and 6 had no effect (Figure 7). This suggests that the sequence GTGCACAGAGGGT, located between nucleotides −208 and −196, is involved in transcription regulation of the ephrin-A4 gene. We have not identified any consensus sequence for a transcription factor in this sequence using the MatInspector program [29].

DISCUSSION

To understand the regulation of the human ephrin-A4 (*EFNA4*) gene, we have characterized the ephrin-A4 promoter in detail to identify proximal *cis* elements involved in transcriptional regulation. To our knowledge, this is the first detailed analysis of a promoter regulating the expression of an ephrin gene. Several reports [24,25] have described promoters for the Eph receptors binding the ephrin ligands. In the present study, the sequence of 679 nucleotides of the upstream-regulatory region of the *EFNA4*

gene is characterized, with putative binding sites for several known transcription factors such as NF-Y, MEF2, c-*myb* and GATA. We cannot exclude that the promoter activity may extend to more distal parts of the studied sequence. To study transcriptional regulation of the ephrin-A4 gene, a reporter-gene construct was generated with the ephrin-A4 promoter in front of the firefly luciferase gene. Transient transfection of this construct into different haemopoietic cell lines revealed that the isolated ephrin-A4 promoter is able to direct transcription and contains *cis* elements necessary for cell-specific gene expression.

To determine the *cis* elements in the ephrin-A4 promoter-regulating transcription, various deletion constructs of the promoter driving a luciferase reporter gene were generated. The 366 nucleotides upstream from the start codon were needed to obtain optimal transcription activity from this promoter. Within these 366 nucleotides, three distinct regions, namely regions I, II and III, contributed to the transcription activity. The elements responsible for the activity, contributed by each of these three regions, were examined both by EMSA and by mutation analyses, to identify the proteins binding to and the critical nucleotides in these promoter elements.

A strong retardation of a ubiquitous nature was observed using a DNA fragment-covering region III in the EMSA. In this region, an inverted CCAAT box was shown to be involved in protein binding defined by competition studies with either mutated CCAAT box or an unrelated CCAAT box from the CD10 promoter. In addition, mutation of the CCAAT box abrogated the ephrin-A4 promoter activity in transfected cells. The CCAAT-box motif is found in 25–30% of all characterized promoters, including housekeeping genes, cell cycle-regulated genes and tissue-specific and developmentally regulated genes [34,36]. The major protein binding to CCAAT boxes is the NF-Y transcription activator [34,36]. A supershift was observed using an NF-Y-specific antibody in the EMSA reaction, indicating the capacity of this factor to bind the ephrin-A4 promoter. In addition, binding of NF-Y to the ephrin-A4 promoter was abrogated after mutation of the inverted CCAAT box. NF-Y is a broadly expressed heterotrimeric transcription factor involved in basal transcription, as well as in both positive and negative regulative responses. The NF-Y trimer is not a strong transcription activator and it may play a role as a general promoter organizer [38]. In addition to being involved in DNA bending [39,40], NF-Y can cooperate with the co-activators p300 (a transcriptional activator protein required to derive p53 expression) [41] and histone acetyl transferases P/CAF (p300/cAMP-response-element-binding protein) and GCN5 (positive general control of transcription-5) [42,43]. Several studies indicate that binding of NF-Y to the CCAAT box also recruits other transcription factors [44,45]. The observation that mutation of the CCAAT box nearly abolishes the reporter-gene activity, directed from the ephrin-A4 promoter, may indicate that NF-Y acts as a promoter organizer, since deletion studies showed that this region contributed only weakly to the overall promoter activity. There are also examples where NF-Y activates transcription through cooperation with other transcription factors [35], such as the combination of NF-Y and Sp1 involved in the cell-cycle regulation of A-myb [46] and in the UV-induced increase in transcription of the MDR gene [47]. A GC box, which is a consensus-binding site for the Sp1/Sp3 transcription factors, is located 110 bp upstream for the inverted CCAAT box in the ephrin-A4 promoter. However, no transcription activity was associated with this GC box.

A putative myb element with a weak consensus sequence [48] was identified 4 nucleotides downstream of the inverted CCAAT box. A similar combination of a CCAAT box and an identical putative myb site has been described in the 5′-UTR of the CD34 gene [37]. Here, the CCAAT box is orientated in the forward direction, whereas in the ephrin-A4 promoter it is inverted. Also, the distance between the two elements is different, with 11 and 4 nucleotides distance respectively. There are other examples where both forward orientated and inverted CCAAT boxes cooperate with another element and with comparable spacing. The endoplasmic reticulum (ER) stress response element (ERSE) consists of a forward orientated CCAAT box binding NF-Y and a binding site for the p50ATF6 transcription factor. Both NF-Y and p50ATF6 must bind DNA to increase transcription and the distance between these two sites is critical [49]. Another ERSE, ERSE-II, consists of the same two binding sites, affecting transcription in the same way as shown for ERSE, but here the CCAAT box is inverted and the spacing between the CCAAT box and the p50ATF6-binding site is different [50].

Mutations introduced in the putative myb site decreased the ephrin-A4 promoter activity when compared with the mutations introduced in the inverted CCAAT box. This observation can be explained by a less-stable NF-Y binding to the promoter since NF-Y was not able to bind a DNA fragment lacking the putative myb site. Close examination of the importance of distinct nucleotides in the sequence surrounding the CCAAT box revealed that binding of NF-Y is dependent on a part of the putative myb site, where G is crucial and also the following TTAs are involved (the three last nucleotides in the putative myb site). In addition, the mutation of the two T nucleotides downstream of the putative myb site also affects NF-Y binding, suggesting that a part of the putative myb site and that of the 3′-sequence involves an important binding site including at least six nucleotides, GTTATT. The AT-enriched stretch that surrounds the CCAAT box was also investigated since it has been shown that AT-enriched stretches can function as a non-consensus TATA box or constitute a binding site for other important activator(s) [51,52]. Close examination of distinct nucleotides involved in NF-Y binding and activity reveal that the AT-enriched stretch by itself is not crucial for NF-Y binding. Thus our conclusion is that stable binding of NF-Y to the CCAAT box is dependent on the integrity of the surrounding sequence.

There are examples of NF-Y involvement in both constitutive gene expression and induced gene expression. In the α1-collagen promoter, NF-Y serves only as a constitutive activator and contacts DNA on both sides of the CCAAT box in a sequence-independent manner [53,54], whereas the expression of the CHOP gene is induced by the ER stress involving NF-Y [49]. In this context, NF-Y contacts the CHOP promoter in a sequence-specific manner close to the CCAAT box (ER stress-inducible 'i' domain). Both specific sequence of the 'i' domain and the region of NF-Y that makes contact with such a domain may form a new contact surface for a putative ER stress-inducible factor [55]. Thus one can speculate that if NF-Y directly interacts with an additional specific sequence in the ephrin-A4 promoter, it may change its morphology followed by the binding of specific factor(s) responsible for both activation and cell-specific regulation capacity of the ephrin-A4 promoter. Although region III in the ephrin-A4 promoter binds NF-Y in many different cell types shown by EMSA, promoter activity correlates with the expression pattern, suggesting that this region might be involved in the cell-specific regulation of the gene. One can also speculate that binding of NF-Y to the CCAAT box is dependent on another transcription factor binding to the specific surrounding sequence. A corresponding example is the increased binding affinity of NF-Y to the CCAAT box in the invariant chain promoter after binding of the SP1 transcription factor to a GC box in this promoter [56].

The region III part of the ephrin-A4 promoter, although absolutely necessary for the transcription activity, is not able to direct high levels of transcription. The two other regions, regions I and II, contribute significantly to the promoter activity in transfectants and were further analysed. Protein binding, as shown by retardation of a probe in the EMSA, was only detected in the region I DNA fragment. By UV cross-linking experiments, a factor of approx. 150 kDa was identified to be binding this region. A consensus-binding site for the MEF2 transcription factors was identified in this region, as well as a consensus c-*myb*-binding site. Several factors are known to bind the MEF2-binding site [57], but to this end we have not identified the specific factor(s) binding this site. No band-shift was observed using a region II DNA fragment in the EMSA, although the initial promoter-reporter gene data indicated the existence of an important regulatory element in this region. A linker scanning analysis was performed on region II, and mutations including nucleotides -208 to -196 decreased the reporter-gene activity in transfected cells supporting the importance of this site in transcription regulation. No obvious putative transcription factor binding this region was identified.

In conclusion, we have characterized the human ephrin-A4 promoter and have suggested that it is able to direct the cell-specific gene expression. Three distinct regions in the promoter were shown to contribute to the transcription regulation of the ephrin-A4 gene in lymphoid cells. In particular, an inverted CCAAT box is essential for transcription activity, which is capable of binding the transcription factor NF-Y.

REFERENCES

1 Pandey, A., Lindberg, R. A. and Dixit, V. M. (1995) Cell signalling. Receptor orphans find a family. Curr. Biol. **5**, 986–989

2 Eph Nomenclature Committee (1997) Unified nomenclature for Eph family receptors and their ligands, the ephrins. Cell (Cambridge, Mass.) **90**, 403–404

3 Bolognese, F., Wasner, M., Dohna, C. L., Gurtner, A., Ronchi, A., Muller, H., Manni, I., Mossner, J., Piaggio, G., Mantovani, R. et al. (1999) The cyclin B2 promoter depends on NF-Y, a trimer whose CCAAT-binding activity is cell-cycle regulated. Oncogene **18**, 1845–1853

4 Kalo, M. S. and Pasquale, E. B. (1999) Signal transfer by Eph receptors. Cell Tissue Res. **298**, 1–9

5 Mellitzer, G., Xu, Q. and Wilkinson, D. G. (2000) Control of cell behaviour by signalling through Eph receptors and ephrins. Curr. Opin. Neurobiol. **10**, 400–408

6 Huai, J. and Drescher, U. (2001) An ephrin A-dependent signaling pathway controls integrin function and is linked to the tyrosine phosphorylation of a 120 kDa protein. J. Biol. Chem. **276**, 6689–6694

7 George, S. E., Simokat, K., Hardin, J. and Chisholm, A. D. (1998) The VAB-1 Eph receptor tyrosine kinase functions in neural and epithelial morphogenesis in C. elegans. Cell (Cambridge, Mass.) **92**, 633–643

8 Caras, I. W. (1997) A link between axon guidance and axon fasciculation suggested by studies of the tyrosine kinase receptor EphA5/REK7 and its ligand ephrin-A5/AL-1. Cell Tissue Res. **290**, 261–264

9 Holder, N., Durbin, L. and Cooke, J. (2000) Eph receptors and ephrins are key regulators of morphogenesis. Paper presented in Ernst Schering Res. Found. Workshop, pp. 123–147

10 Xu, Q., Alldus, G., Macdonald, R., Wilkinson, D. G. and Holder, N. (1996) Function of the Eph-related kinase rtk1 in patterning of the zebrafish forebrain. Nature (London) **381**, 319–322

11 Winning, R. S., Scales, J. B. and Sargent, T. D. (1996) Disruption of cell adhesion in Xenopus embryos by Pagliaccio, an Eph-class receptor tyrosine kinase. Dev. Biol. **179**, 309–319

12 Pandey, A., Shao, H., Marks, R. M., Polverini, P. J. and Dixit, V. M. (1995) Role of B61, the ligand for the Eck receptor tyrosine kinase, in TNF-α-induced angiogenesis. Science **268**, 567–569

13 Wang, H. U., Chen, Z. F. and Anderson, D. J. (1998) Molecular distinction and angiogenic interaction between embryonic arteries and veins revealed by ephrin-B2 and its receptor Eph-B4. Cell (Cambridge, Mass.) **93**, 741–753

14 Dodelet, V. C. and Pasquale, E. B. (2000) Eph receptors and ephrin ligands: embryogenesis to tumorigenesis. Oncogene **19**, 5614–5619

15 Kozlosky, C. J., Maraskovsky, E., McGrew, J. T., VandenBos, T., Teepe, M., Lyman, S. D., Srinivasan, S., Fletcher, F. A., Gayle, R. B. 3rd, and Cerretti, D. P. (1995) Ligands for the receptor tyrosine kinases hek and elk: isolation of cDNAs encoding a family of proteins. Oncogene **10**, 299–306

16 Aasheim, H. C., Munthe, E., Funderud, S., Smeland, E. B., Beiske, K. and Logtenberg, T. (2000) A splice variant of human ephrin-A4 encodes a soluble molecule that is secreted by activated human B lymphocytes. Blood **95**, 221–230

17 Cerretti, D. P., Lyman, S. D., Kozlosky, C. J., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Valentine, V., Kirstein, M. N., Shapiro, D. N. and Morris, S. W. (1996) The genes encoding the Eph-related receptor tyrosine kinase ligands LERK-1 (EPLG1, Epl1), LERK-3 (EPLG3, Epl3), and LERK-4 (EPLG4, Epl4) are clustered on human chromosome 1 and mouse chromosome 3. Genomics **33**, 277–282

18 Gale, N. W., Holland, S. J., Valenzuela, D. M., Flenniken, A., Pan, L., Ryan, T. E., Henkemeyer, M., Strebhardt, K., Hirai, H., Wilkinson, D. G. et al. (1996) Eph receptors and ligands comprise two major specificity subclasses and are reciprocally compartmentalized during embryogenesis. Neuron **17**, 9–19

19 Janis, L. S., Cassidy, R. M. and Kromer, L. F. (1999) Ephrin-A binding and EphA receptor expression delineate the matrix compartment of the striatum. J. Neurosci. **19**, 4962–4971

20 Choi, S., Jeong, J., Kim, T. and Park, S. (1999) Characterization of ephrin-A1 and ephrin-A4 as ligands for the EphA8 receptor protein tyrosine kinase. Mol. Cells **9**, 440–445

21 Theil, T., Frain, M., Gilardi-Hebenstreit, P., Flenniken, A., Charnay, P. and Wilkinson, D. G. (1998) Segmental expression of the EphA4 (Sek-1) receptor tyrosine kinase in the hindbrain is under direct transcriptional control of Krox-20. Development **125**, 443–452

22 Chen, J. and Ruley, H. E. (1998) An enhancer element in the EphA2 (Eck) gene sufficient for rhombomere-specific expression is activated by HOXA1 and HOXB1 homeobox proteins. J. Biol. Chem. **273**, 24670–24675

23 Taneja, R., Thisse, B., Rijli, F. M., Thisse, C., Bouillet, P., Dolle, P. and Chambon, P. (1996) The expression pattern of the mouse receptor tyrosine kinase gene MDK1 is conserved through evolution and requires Hoxa-2 for rhombomere-specific expression in mouse embryos. Dev. Biol. **177**, 397–412

24 Jeong, J., Choi, S., Gu, C., Lee, H. and Park, S. (2000) Genomic structure and promoter analysis of the mouse EphA8 receptor tyrosine kinase gene. DNA Cell Biol. **19**, 291–300

25 Dottori, M., Down, M., Huttmann, A., Fitzpatrick, D. R. and Boyd, A. W. (1999) Cloning and characterization of EphA3 (Hek) gene promoter: DNA methylation regulates expression in hematopoietic tumor cells. Blood **94**, 2477–2486

26 Logan, C., Wizenmann, A., Drescher, U., Monschau, B., Bonhoeffer, F. and Lumsden, A. (1996) Rostral optic tectum acquires caudal characteristics following ectopic engrailed expression. Curr. Biol. **6**, 1006–1014

27 Shigetani, Y., Funahashi, J. I. and Nakamura, H. (1997) En-2 regulates the expression of the ligands for Eph type tyrosine kinases in chick embryonic tectum. Neurosci. Res. **27**, 211–217

28 Rhinn, M., Dierich, A., Le Meur, M. and Ang, S. (1999) Cell autonomous and non-cell autonomous functions of Otx2 in patterning the rostral brain. Development **126**, 4295–4304

29 Quandt, K., Frech, K., Karas, H., Wingender, E. and Werner, T. (1995) MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res. **23**, 4878–4884

30 Schreiber, E., Matthias, P., Muller, M. M. and Schaffner, W. (1989) Rapid detection of octamer binding proteins with 'mini-extracts' prepared from a small number of cells. Nucleic Acids Res. **17**, 6419

31 Ishimaru, F., Mari, B. and Shipp, M. A. (1997) The type 2 CD10/neutral endopeptidase 24.11 promoter: functional characterization and tissue-specific regulation by CBF/NF-Y isoforms. Blood **89**, 4136–4145

32 Swanson, B. J., Jack, H. M. and Lyons, G. E. (1998) Characterization of myocyte enhancer factor 2 (MEF2) expression in B and T cells: MEF2C is a B cell-restricted transcription factor in lymphocytes. Mol. Immunol. **35**, 445–458

33 Wingender, E., Chen, X., Hehl, R., Karas, H., Liebich, I., Matys, V., Meinhardt, T., Pruss, M., Reuter, I. and Schacherer, F. (2000) TRANSFAC: an integrated system for gene expression regulation. Nucleic Acids Res. **28**, 316–319

34 Mantovani, R. (1998) A survey of 178 NF-Y binding CCAAT boxes. Nucleic Acids Res. **26**, 1135–1143

35 Mantovani, R. (1999) The molecular biology of the CCAAT-binding factor NF-Y. Gene **239**, 15–27

36 Bucher, P. (1990) Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. J. Mol. Biol. **212**, 563–578

37 Radomska, H. S., Satterthwaite, A. B., Taranenko, N., Narravula, S., Krause, D. S. and Tenen, D. G. (1999) A nuclear factor Y (NFY) site positively regulates the human CD34 stem cell gene. Blood **94**, 3772–3780

38 de Silvio, A., Imbriano, C. and Mantovani, R. (1999) Dissection of the NF-Y transcriptional activation potential. Nucleic Acids Res. **27**, 2578–2584

39 Liberati, C., Ronchi, A., Lievens, P., Ottolenghi, S. and Mantovani, R. (1998) NF-Y organizes the γ-globin CCAAT boxes region. J. Biol. Chem. **273**, 16880–16889

40 Ronchi, A., Bellorini, M., Mongelli, N. and Mantovani, R. (1995) CCAAT-box binding protein NF-Y (CBF, CP1) recognizes the minor groove and distorts DNA. Nucleic Acids Res. **23**, 4565–4572

41 Faniello, M. C., Bevilacqua, M. A., Condorelli, G., de Crombrugghe, B., Maity, S. N., Avvedimento, V. E., Cimino, F. and Costanzo, F. (1999) The B subunit of the CAAT-binding factor NFY binds the central segment of the co-activator p300. J. Biol. Chem. **274**, 7623–7626

42 Jin, S. and Scotto, K. W. (1998) Transcriptional regulation of the MDR1 gene by histone acetyltransferase and deacetylase is mediated by NF-Y. Mol. Cell Biol. **18**, 4377–4384

43 Currie, R. A. (1998) NF-Y is associated with the histone acetyltransferases GCN5 and P/CAF. J. Biol. Chem. **273**, 1430–1434

44 Wright, K. L., Vilen, B. J., Itoh-Lindstrom, Y., Moore, T. L., Li, G., Criscitiello, M., Cogswell, P., Clarke, J. B. and Ting, J. P. (1994) CCAAT box binding protein NF-Y facilitates *in vivo* recruitment of upstream DNA binding transcription factors. EMBO J. **13**, 4042–4053

45 Linhoff, M. W., Wright, K. L. and Ting, J. P. (1997) CCAAT-binding factor NF-Y and RFX are required for *in vivo* assembly of a nucleoprotein complex that spans 250 base pairs : the invariant chain promoter as a model. Mol. Cell Biol. **17**, 4589–4596

46 Facchinetti, V., Lopa, R., Spreafico, F., Bolognese, F., Mantovani, R., Tavner, F., Watson, R., Introna, M. and Golay, J. (2000) Isolation and characterization of the human A-*myb* promoter : regulation by NF-Y and Sp1. Oncogene **19**, 3931–3940

47 Hu, Z., Jin, S. and Scotto, K. W. (2001) Transcriptional activation of the MDR1 gene by UV irradiation. Role of NF-Y and Sp1. J. Biol. Chem. **275**, 2979–2985

48 Ording, E., Bergholtz, S., Brendeford, E. M., Jamin, N. and Gabrielsen, O. S. (1996) Flexibility in the second half-site sequence recognised by the c-Myb R2 domain – *in vitro* and *in vivo* analysis. Oncogene **13**, 1043–1051

49 Yoshida, H., Okada, T., Haze, K., Yanagi, H., Yura, T., Negishi, M. and Mori, K. (2000) ATF6 activated by proteolysis binds in the presence of NF-Y (CBF) directly to the *cis*-acting element responsible for the mammalian unfolded protein response. Mol. Cell Biol. **20**, 6755–6767

50 Kokame, K., Kato, H. and Miyata, T. (2001) Identification of ERSE-II, a new *cis*-acting element responsible for the ATF6-dependent mammalian unfolded protein response. J. Biol. Chem. **276**, 9199–9205

51 Fuller, K. and Storb, U. (1997) Identification and characterization of the murine Rag1 promoter. Mol. Immunol. **34**, 939–954

52 Currie, R. A. (1997) Functional interaction between the DNA binding subunit trimerization domain of NF-Y and the high mobility group protein HMG-I(Y). J. Biol. Chem. **272**, 30880–30888

53 Bi, W., Wu, L., Coustry, F., de Crombrugghe, B. and Maity, S. N. (1997) DNA binding specificity of the CCAAT-binding factor CBF/NF-Y. J. Biol. Chem. **272**, 26562–26572

54 Liang, S. G. and Maity, S. N. (1998) Pathway of complex formation between DNA and three subunits of CBF/NF-Y. Photocross-linking analysis of DNA–protein interaction and characterization of equilibrium steps of subunit interaction and DNA binding. J. Biol. Chem. **273**, 31590–31598

55 Ubeda, M. and Habener, J. F. (2000) CHOP gene expression in response to endoplasmic-reticular stress requires NFY interaction with different domains of a conserved DNA-binding element. Nucleic Acids Res. **28**, 4987–4997

56 Wright, K. L., Moore, T. L., Vilen, B. J., Brown, A. M. and Ting, J. P. (1995) Major histocompatibility complex class II-associated invariant chain gene expression is up-regulated by cooperative interactions of Sp1 and NF-Y. J. Biol. Chem. **270**, 20978–20986

57 Fickett, J. W. (1996) Quantitative discrimination of MEF2 sites. Mol. Cell Biol. **16**, 437–441