The EMBO Journal vol.10 no.11 pp.3409-3417, 1991

A novel lineage-specific nuclear factor regulates *mb-1* gene transcription at the early stages of B cell differentiation

James Hagman, Adam Travis and Rudolf Grosschedl

Howard Hughes Medical Institute, Departments of Microbiology and Immunology, and Biochemistry and Biophysics, University of California, San Francisco, CA 94143-0414, USA

Communicated by R.A.Flavell

The *mb-1* gene, which encodes a protein associated with membrane-bound antibody, is expressed only at the early stages of B cell differentiation. To gain insight into the mechanisms that underlie temporally regulated gene expression, we examined the *mb-1* promoter region for interactions with cell type-specific DNA binding proteins. Here, we report the characterization of a novel nuclear factor that recognizes the *mb-1* promoter. This DNA binding activity, termed Early B cell Factor, or EBF, is expressed in early stage B cells, but not in late stage B cells, T cells or non-lymphoid cells. EBF recognizes the nucleotide sequence 5'-CAAGGGAAT-3' in the *mb-1* and major histocompatibility complex (MHC) class II A_α^d promoters. The binding of EBF to DNA was characterized by DNase I footprinting and by methylation interference analysis which indicated both major and minor groove contacts. The specificity of EBF binding is distinct from that of other nuclear factors expressed in hematopoietic cells. EBF appears to consist of at least two polypeptides of ~70-75 kDa and 80-85 kDa. The EBF binding site was important for maximal *mb-1* promoter activity in early stage B cells. Moreover, the EBF binding site conferred correct lineage- and stage-specific transcriptional activity upon a heterologous promoter in a context-dependent manner. Thus, EBF appears to represent an important transcriptional regulator of B cell specific gene expression.

Key words: B cell differentiation/DNA binding factor/gene regulation/*mb-1* gene expression

Introduction

The differentiation of B cells from progenitor cells to antibody-secreting plasma cells progresses through a series of developmental stages that are defined by changes in the expression of immunoglobulin (Ig) heavy (H) and light (L) chain genes. The early stages of the B cell lineage are represented by pro-B cells, which do not express complete μ H chains, pre-B cells, which express rearranged μ H chain genes, and finally the more mature B cells, which in addition to H chains express rearranged L chain genes and display IgM molecules on the B cell surface as membrane-bound antigen receptors. IgM-positive resting B cells can be triggered by antigen and other external signals to terminally differentiate into plasma cells which no longer display cell surface membrane-bound IgM.

Membrane-bound IgM is non-covalently associated with other polypeptides (reviewed by Cambier, 1991), including a 32-34 kDa glycoprotein encoded by the *mb-1* gene (Sakaguchi *et al.*, 1988). The amino acid sequence of the *mb-1* gene product, variously termed MB-1 (Sakaguchi *et al.*, 1988), B34 (Hombach *et al.*, 1988), pp32 (Campbell and Cambier, 1990), or IgMα (Hombach *et al.*, 1990), is structurally related to the T cell receptor-associated membrane proteins CD3γ and CD3δ (Sakaguchi *et al.*, 1988). Together, these observations suggest that the *mb-1* encoded polypeptide may be part of a transmembrane signalling structure involved in antigen receptor-mediated B cell activation (Sakaguchi *et al.*, 1988; Reth, 1989). Moreover, the MB-1 protein appears to be required for transport of the IgM complex to the B cell surface, because plasma cells, which do not express endogenous MB-1 molecules, displayed surface-bound IgM only when a transfected *mb-1* gene was co-expressed (Hombach *et al.*, 1990).

Transcripts of the *mb-1* gene are detected only in the early stages of the B cell lineage, including pro-B, pre-B and surface IgM-positive B cells (Sakaguchi *et al.*, 1988). Terminally differentiated plasma cells lack *mb-1* transcripts, coinciding with the disappearance of membrane-bound Ig molecules from the cell surface. This expression pattern establishes *mb-1* as a member of a set of genes that is transcriptionally active at the early, but not late stages of B cell differentiation. This pattern of expression is shared in part by murine MHC class II genes, which are expressed in surface Ig-positive B cells, but not in pre-B or plasma cells (reviewed in Benoist and Mathis, 1990). Another set of genes, including λ5 (Sakaguchi and Melchers, 1986; Sakaguchi *et al.*, 1986). Vpre-B (Kudo and Melchers, 1987) and terminal deoxynucleotidyl transferase (TdT) (Landau *et al.*, 1984) are expressed specifically at the pre-B cell stage. Together, these genes are expressed in a common lineage, but are differentially regulated during B cell development.

Studies aimed at defining mechanisms which underlie the temporal regulation of B lineage-specific genes have focused to a large extent on the Ig κ gene, which is transcriptionally activated in mature B cells and is expressed throughout the remaining stages of differentiation (reviewed by Staudt and Lenardo, 1991). Two transcriptional regulators, NFκB and Oct-2, have been implied to be crucial for the temporal regulation of this gene. NFκB, which binds to the κ light chain intron enhancer, is undetectable in the nuclei of pre-B cells and its appearance in the nucleus in response to external cues or signals parallels the up-regulation of κ expression (reviewed by Lenardo and Baltimore, 1989). The lymphoid-specific homeodomain protein Oct-2 binds an octamer sequence that is important for κ promoter function in B lineage cells (reviewed by Staudt and Lenardo, 1991). The *oct-2* gene is expressed at low levels in pre-B cells and is up-regulated in response to signals which also activate κ expression (Staudt *et al.*, 1988).

© Oxford University Press

In contrast, the regulation of lymphoid-specific genes that are first activated and then shut off at defined stages of cell differentiation is not understood. We have selected the upstream control region of the *mb-1* gene as a model system for studying mechanisms that govern temporally regulated gene expression in the B cell lineage. The *mb-1* promoter lacks an identifiable TATA element and directs transcription initiation from multiple start sites (Travis *et al.*, 1991b). The activity of an *mb-1* promoter fragment in transfection assays suggested that positive regulatory factors specifically activate *mb-1* transcription in the early stages of B cell differentiation. Alternatively, the promoter may be negatively regulated in other cell types. To this end, two DNA-binding factors, MUF2 and Sp1, were identified which recognize functionally important proximal promoter region sequences *in vitro*. However, the presence of these factors in cells which do not express *mb-1* transcripts suggests that additional cell type-specific factors regulate the specificity of *mb-1* expression.

Here we report the description of a novel DNA binding activity termed Early B cell Factor, or EBF, which recognizes a specific nucleotide sequence in the distal region of the *mb-1* promoter. EBF can be detected only in nuclear extracts of cells that express *mb-1*, and not in extracts of other lymphoid or non-lymphoid cell types. Analysis of EBF binding to the *mb-1* promoter and characterization of the protein–DNA complex indicated that this factor is distinct from BSAP, a previously identified transcriptional activator which shares the same expression pattern (Barberis *et al.*, 1990). The EBF binding site contributes to *mb-1* promoter function and, together with flanking sequences, can confer early B cell-specific expression upon a heterologous gene.

Results

Specific binding of EBF to the distal promoter region of the mb-1 gene

To identify and localize binding sites for nuclear factors in the *mb-1* promoter, we incubated a labeled DNA fragment comprising the distal promoter region (−252/−113) with nuclear extracts, and digested the products with DNase I. Nucleotides extending from −160 to −179 on the coding, and −164 to −181 on the non-coding strands were protected from DNase I digestion by nuclear extract from the pre-B cell line PD36 (Figure 1A, lanes 4 and 9). A few nucleotides flanking the protected region displayed an enhanced sensitivity of DNase I cleavage. In contrast, no protection was observed when the DNA probe was incubated with nuclear extract from the fibroblast cell line, NIH3T3 (Figure 1A, lanes 5 and 10). Likewise, no protection was observed with nuclear extracts from cell lines representing plasma cells and T cells (data not shown). Therefore, these data localize a binding site for a putative B cell-specific nuclear factor to the distal region of the *mb-1* promoter (Figure 1B).

The presence of this nuclear factor in various cell types was further examined by analyzing nuclear proteins from a panel of lymphoid and non-lymphoid cell lines in an electrophoretic mobility shift assay (Singh *et al.*, 1986). Incubation of a labelled distal promoter region fragment (−252/−158) with nuclear extracts from pro-B, pre-B and surface Ig-positive B cells resulted in the formation of a single major protein–DNA complex (Figure 2, top panel). In contrast, nuclear extracts derived from myelomas repre-

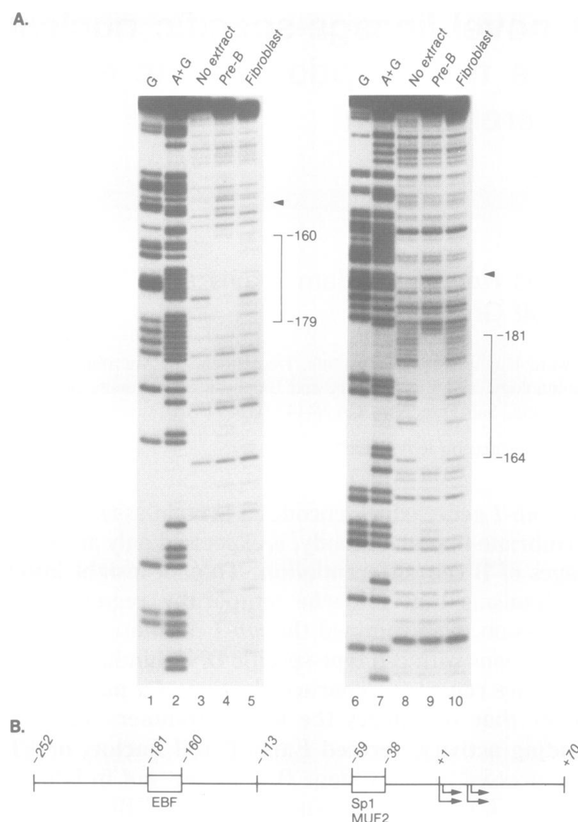

Fig. 1. DNase I footprint analysis of the distal *mb-1* promoter region. (**A**) The coding (lanes 1–5) or non-coding (lanes 6–10) strands of a distal *mb-1* promoter probe (−252/−113) were 5′ end-labeled with [γ-^{32}P]ATP, incubated with nuclear extract as indicated, digested with DNase I and separated on denaturing gels. Lanes 1, 6 and 2, 7 represent G and A+G Maxam–Gilbert sequencing ladders respectively (Maniatis *et al.*, 1982). Brackets indicate the extent of protection from DNase I cleavage. The position of nucleotides are assigned relative to the 5′ most major site of transcription initiation (+1). Arrowheads indicate enhanced DNase I cleavage sites. (**B**) Relative location of binding sites for nuclear factors in the *mb-1* promoter. The approximate boundaries of the DNase I protected regions are represented by boxes. Below the boxes, the nuclear factors that bind the DNase I protected regions in electrophoretic mobility shift assays are indicated. Major sites of transcription initiation are designated as arrows (Travis *et al.*, 1991b).

senting late stage B cells, T lymphoid, myeloid and other non-lymphoid cell lines did not contain this complex. The DNA binding activity detected in murine early B cell lines was also observed in nuclear extracts from a human B cell line (BJAB) and adult mouse splenic B cells (data not shown), indicating that both human and normal mouse cells contain a factor with similar DNA binding specificity. Minor complexes with slightly faster mobility were observed in two cell lines which do not express *mb-1* (SP2 and NIH3T3). The relationship of these complexes to EBF is unknown. To confirm that the absence of this binding activity from late stage B cell and non-B cell extracts is not an artefact of sample preparation, we examined all nuclear extracts for the presence of the 'ubiquitous' Oct-1 protein (Figure 2, bottom panel). Similar levels of Oct-1 were detected in all of the analyzed nuclear extracts. Taken together, these data confirm that the distal *mb-1* promoter region is recognized by an early B cell-specific DNA binding factor with a cell type distribution paralleling the expression of the endogenous *mb-1* gene. Therefore, we have termed this DNA binding activity Early B cell Factor (EBF).

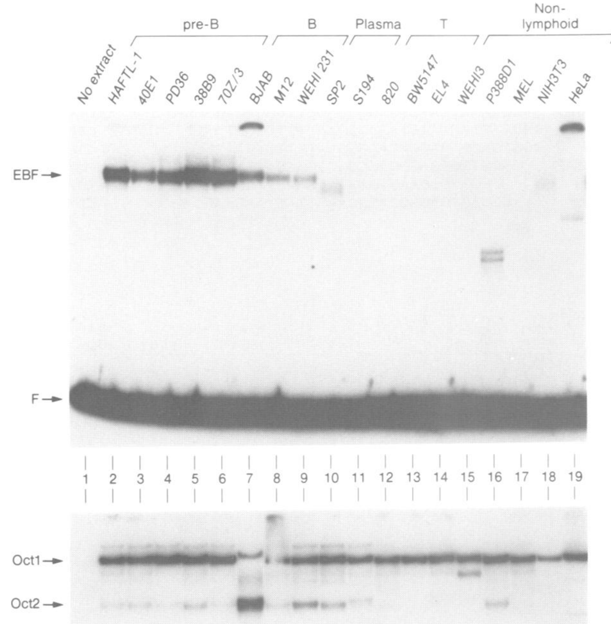

Fig. 2. Cell type-specific distribution of EBF. 2.5 μg of crude nuclear extract prepared from various lymphoid and non-lymphoid cell lines were analyzed in an electrophoretic mobility shift assay with ^{32}P-labeled DNA probes. (**Top**) Identification of EBF with an *mb-1* promoter fragment (−252/−158) comprising the EBF binding site and flanking sequences. (**Bottom**) Detection of Oct-1 with a Vκ41 octamer probe (sequence is shown in Figure 4C). The positions of EBF, free probe (F) and the octamer binding proteins Oct-1 and Oct-2 are indicated.

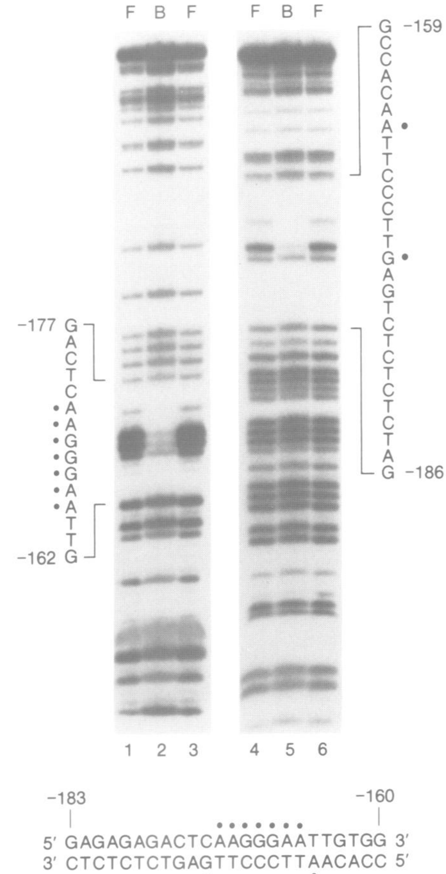

Fig. 3. Methylation interference analysis of the EBF−DNA complex. The coding (lanes 1−3) or non-coding (lanes 4−6) strands of the *mb-1* promoter probe (−243/−126) were 3′ end-labeled with ^{32}P, partially methylated with dimethylsulfate (DMS) and incubated with pre-B cell (PD36) nuclear extract. Bound (B) and free (F) probe were separated in an electrophoretic mobility shift assay and processed as described in Materials and methods. Although the conditions used for strand scission favored the cleavage of methylated guanine residues, the detection of adenines was enhanced in the region of the EBF binding site. Bases which interfere with EBF binding when methylated are indicated by filled circles within the intervals shown. The *mb-1* promoter sequence in the region of the EBF binding site is shown below with a summary of the interference pattern.

Sequence specificity of EBF binding

To determine the contacts between EBF and its binding site in the *mb-1* promoter, we performed a methylation interference assay with a partially methylated DNA probe (Figure 3). Methylation at seven contiguous G and A residues on the coding strand strongly interfered with the binding of EBF, suggesting that EBF contacts both the major groove (inhibited by G-N^7 methylation) and minor groove (inhibited by A-N^3 methylation) of the binding site (Siebenlist and Gilbert, 1980). On the non-coding strand, EBF contacts one G and one A residue which flank the stretch of purines contacted on the coding strand. Thus, interference with EBF binding was observed over nine contiguous nucleotides that appear to represent a core binding site within the region protected from DNase I digestion.

The sequence specificity of EBF binding was further demonstrated by competitive binding assays with wild-type and mutated oligonucleotides (Figure 4A). EBF binding to a distal *mb-1* promoter probe was efficiently competed by a 30 bp wild-type oligonucleotide containing the entire DNase I protected region (−183/−154; Figure 4B, lanes 3−5). The oligonucleotides m3, m4 and LS (−177/−171) contain mutations in the EBF core binding site and failed to compete significantly for EBF binding (lanes 6−8 and 15−20). A single point mutation that changes the C residue at nucleotide −173 to an A residue (oligonucleotide m2; lanes 12−14) decreased competition for EBF binding more than 30-fold. In contrast, the mutated oligonucleotide m5 competed as well as the wild-type oligonucleotide (lanes 21−23), and the mutated m1 oligonucleotide competed with 4-fold higher efficiency than the wild-type sequence (lanes 9−11). The increased efficiency of competition with m1 suggests that sequences flanking the core binding site (as defined by methylation interference analysis) influence EBF binding. Together, these data indicate that EBF recognizes the core sequence 5′-CAAGGGAAT-3′ and additional flanking nucleotides that may contribute to binding.

EBF recognizes an MHC class II promoter sequence

A survey of known regulatory regions of genes that are expressed in early B cells revealed two candidate EBF recognition sites in the promoters of MHC class II A_α chain genes (Figure 4C). A sequence which is identical at all nine nucleotide positions with the EBF core binding site in the *mb-1* promoter is present in the A_α^d promoter region between the proximal X and Y boxes (−60/−52; Dedrick and Jones, 1990). An oligonucleotide that contains this sequence competed for EBF binding with only 3-fold lower efficiency than the cognate EBF binding site of the *mb-1* promoter (Figure 4D, compare lanes 3−5 and 12−14). This suggests that the A_α^d promoter may be regulated by EBF. We have also identified a related sequence between the proximal X and Y boxes of the MHC class II A_α^k pro-

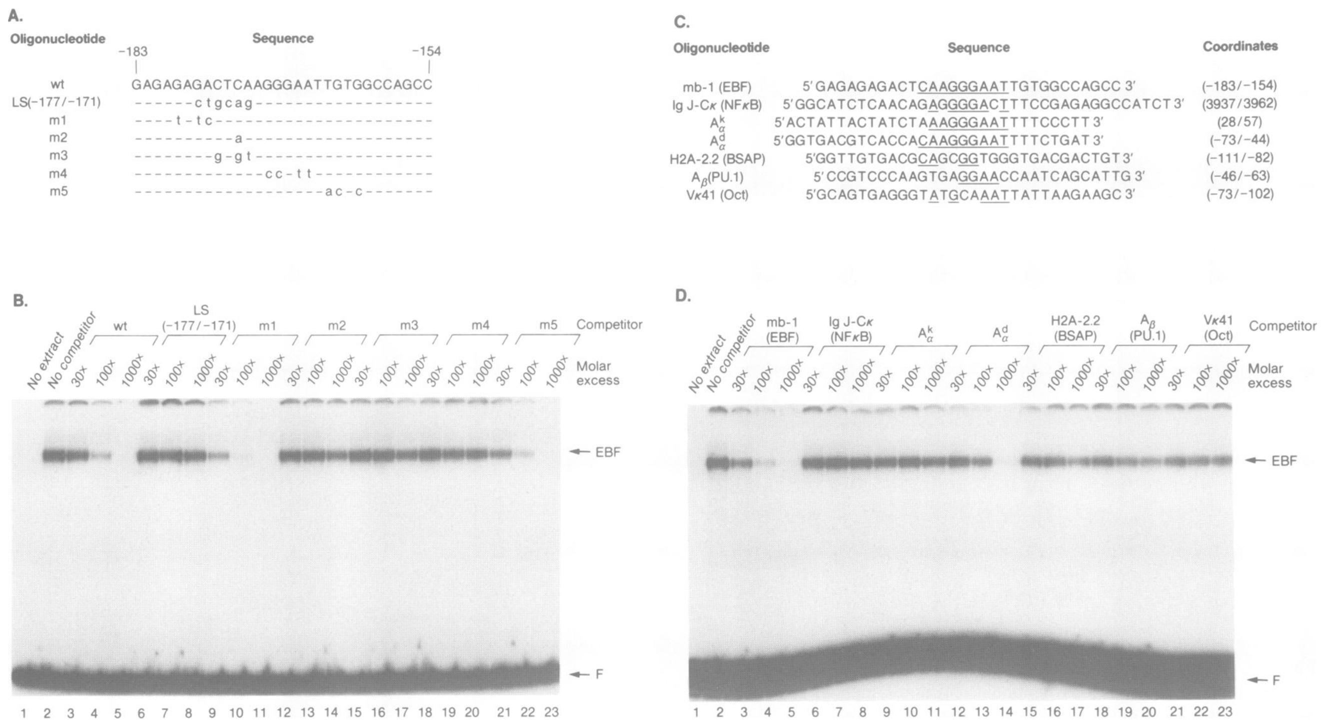

Fig. 4. The DNA binding specificity of EBF and its relationship to other known nuclear factors that regulate B cell-specific gene expression. (**A**) Sequences of synthetic oligonucleotides used as competitors in electrophoretic mobility shift assays. The *mb-1* wild-type (wt) sequence is indicated as upper case letters or as a dashed line and mutations are shown in lower case letters. Clustered mutations introduced in the oligonucleotide LS(−177/−171) are identical with mutations in the promoter sequence shown in Figure 6A. (**B**) Electrophoretic mobility shift assay with a ^{32}P-labeled *mb-1* promoter fragment (−252/−158). Labeled DNA probe and excess unlabeled competitors were incubated with 1 μg pre-B cell (38B9) nuclear extract and separated by native PAGE. The position of the EBF−DNA complex (EBF) and free probe (F) are indicated. (**C**) Sequences of synthetic oligonucleotides used as competitors in electrophoretic mobility shift assays. Nuclear factors that have been reported to bind the sequences *in vitro* are indicated in parentheses. Nucleotide sequences of the oligonucleotides are arranged for optimal alignment with the EBF core binding sequence of the *mb-1* promoter; nucleotides which match the core sequence are underlined. References for sequences and reported coordinates in transcriptional control regions are: Ig J-Cκ (NFκB): murine Ig κ major intron (Max *et al.*, 1981); $A_\alpha{}^k$: murine MHC class II $A_\alpha{}^k$ promoter (Gravallese *et al.*, 1989); murine MHC class II $A_\alpha{}^d$ promoter (Dedrick and Jones, 1990); H2A-2.2 (BSAP): sea urchin late histone H2A-2.2 promoter (Barberis *et al.*, 1990); A_β (PU.1): non-coding strand sequence from murine MHC class II A_β promoter (Celada *et al.*, 1988); Ig Vκ41 (octamer): murine Vκ41 promoter (Bergman *et al.*, 1984). (**D**) Electrophoretic mobility shift assay conducted as described in (B), with competitors shown in (C).

moter (5′-CAAGAGAAT-3′; Gravellese *et al.*, 1989), but have not tested whether this sequence binds EBF.

A second candidate EBF site was identified within the upstream region of the MHC class II $A_\alpha{}^k$ gene (located near −1130; Figure 4C). This lipopolysaccharide (LPS) response element, or LRE, has been reported to bind a nuclear factor which increases in abundance in LPS-stimulated B cells (Gravallese *et al.*, 1989). Although the LRE is identical at 8/9 positions with the EBF core binding site (5′-AAAG-GGAAT-3′) and has additional matches with flanking nucleotides, this sequence did not compete for EBF binding (Figure 4D, lanes 9−11). This is consistent with the observed lack of competition with the m2 mutant oligonucleotide, which also has an A residue in place of the C residue at the first position of the core EBF binding sequence.

EBF is distinct from BSAP and other nuclear factors

The specific expression of EBF in early stage B cell lines parallels the reported expression pattern of the B cell-specific activator protein (BSAP) (Barberis *et al.*, 1990). BSAP was identified as a mammalian nuclear factor that recognizes a loose consensus sequence present in sea urchin late histone promoters. However, no binding site for BSAP has yet been identified in a B cell-specific gene. To determine whether EBF and BSAP have related binding specificities, we performed a competitive binding assay with the *mb-1* promoter probe and a competitor oligonucleotide containing the BSAP binding site of the sea urchin H2A-2.2 promoter. The BSAP binding site failed to compete with the *mb-1* promoter probe for EBF binding (Figure 4D, lanes 15−17). Conversely, the wild-type *mb-1* oligonucleotide did not compete with a labeled sea urchin late histone promoter probe for BSAP binding (not shown). Together, these data suggest that EBF and BSAP have different DNA-binding specificities.

We extended our analysis of the relationship of EBF to other nuclear factors to include additional DNA-binding proteins that regulate gene expression in B cells. The recognition sites of NFκB and the myeloid/B cell-specific factor PU.1 are purine-rich and show some similarity with the core EBF binding site when aligned with the *mb-1* promoter sequence (Figure 4C). Oligonucleotides containing binding sites for either NFκB (Figure 4D, lanes 6−8), or PU.1 (Figure 4D, lanes 18−20), did not compete with the *mb-1* promoter probe for EBF binding. Likewise, an oligonucleotide containing the recognition site for the B cell-specific nuclear factor Oct-2 did not compete for EBF binding (Figure 4D, lanes 21−23). By these criteria, EBF appears to be a novel B lineage-specific DNA binding factor.

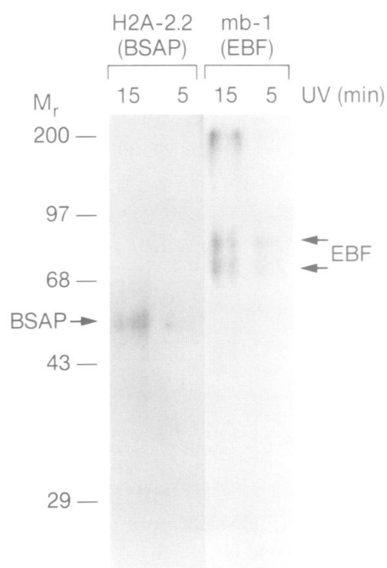

Fig. 5. UV cross-linking of the EBF–DNA complex suggests that EBF is composed of at least two polypeptides. Internally ^{32}P-labeled and 5-BrdU-substituted oligonucleotide probes comprising binding sites for BSAP or EBF were incubated with pre-B cell (PD36) nuclear extract, electrophoresed through a native polyacrylamide gel and irradiated *in situ* with UV light (300 nm) for 5 or 15 min as indicated. EBF–DNA or BSAP–DNA complexes were excised and electrophoresed through an SDS–polyacrylamide gel (Laemmli, 1970) and the cross-linked polypeptides were visualized by autoradiography. The migration and apparent molecular mass (M_r) in kDa of Coomassie blue-stained protein standards are shown for reference. Arrows indicate specifically cross-linked polypeptides.

Polypeptide components of the EBF–DNA complex

To estimate the molecular mass of polypeptides that contact DNA in the EBF–DNA complex, a photoaffinity cross-linking experiment was performed with a uniformly radiolabeled and 5-bromo-deoxyuridine substituted *mb-1* promoter probe. EBF complexes bound to this probe were separated from free probe DNA by electrophoresis through a native polyacrylamide gel. The gel was irradiated *in situ* with UV light to cross-link covalently EBF polypeptides which contact DNA to the oligonucleotide probe. Subsequent separation of DNA–protein adducts by SDS–PAGE (Figure 5) indicated that two polypeptides of 80–85 kDa and 70–75 kDa were cross-linked to the labelled oligonucleotide with nearly equal efficiency. An additional band of >160 kDa, which could contain both polypeptides, was observed when irradiation was continued for longer times. For comparison, cross-linking of BSAP to a sea urchin late histone promoter (H2A-2.2) DNA probe resulted in the expected single band of 50–55 kDa [our estimate is slightly larger than that of Barberis *et al.* (1990), possibly due to the digestion of BSAP–DNA adducts with DNase I prior to SDS–PAGE in the former study]. These data suggest that the EBF–DNA complex contains at least two polypeptides which are distinct from the BSAP polypeptide and bind the *mb-1* promoter as components of a multimeric complex.

A functional EBF binding site is required for high level mb-1 promoter activity in early B cells

To examine the role of the EBF binding site for the transcriptional regulation of the *mb-1* gene, we introduced clustered mutations at this site in the context of the intact *mb-1* promoter. When linked to a heterologous β-globin reporter gene, the 322 bp promoter fragment (−252/+70) can reproduce the cell type-specific pattern of transcription initiation from multiple sites observed for the endogenous *mb-1* promoter (Travis *et al.*, 1991b). Plasmids containing the wild-type or mutant *mb-1* promoter were transfected into the pre-B cell line PD36 and specific transcripts were analyzed by an S1 nuclease protection assay. Transfection efficiency was monitored by measuring the expression of a histone H4 reference gene present on both plasmids (Grosschedl and Baltimore, 1985). Specific transcription was directly quantified using a Molecular Dynamics Phosphor Imager system. As expected, the wild-type *mb-1* promoter directed the synthesis of transcripts from multiple initiation sites (Figure 6B, lane 1). The mutations introduced in the EBF binding site in LS(−177/−171), which completely abrogated EBF binding *in vitro* (Figure 4A and B), reduced the number of specific transcripts 3-fold relative to the wild-type promoter, but did not affect the pattern of transcription initiation (Figure 6B, lane 2). This result indicates that the EBF binding site is important for maximal *mb-1* promoter activity.

The EBF binding site confers early B cell-specific activity upon a heterologous promoter in a context-dependent manner

We generated a second series of plasmids to examine whether the EBF binding site of the distal *mb-1* promoter region can confer early B cell specificity upon a heterologous promoter. One or two copies of a 105 bp fragment (−217/−113) containing the wild-type or a mutated EBF binding site derived from the LS(−177/−171) plasmid were inserted at position −56 of the c-*fos* promoter–chloramphenicol acetyltransferase (CAT) reporter plasmid fosCAT (Gilman *et al.*, 1986). Following transient transfection of the reporter genes together with an RSV-luciferase control plasmid into various cell types, c-*fos* promoter activity was analyzed by measuring CAT levels after the extracts were normalized for luciferase activity (Figure 7). One or two copies of the wild-type distal *mb-1* promoter fragment increased CAT levels in a pre-B cell line (PD36) 4- and 10-fold, respectively. This enhancement was independent of the orientation of the *mb-1* promoter fragments (not shown). In contrast, the distal *mb-1* promoter fragment containing a mutated EBF binding site did not increase CAT levels above that obtained with the fosCAT gene alone. Transfection of the plasmids into M12 lymphoma cells, which contain less EBF than does the PD36 cell line (Figure 2, lane 8 versus 4), resulted in 2- and 4-fold enhancement of CAT activity with one or two copies of the distal promoter sequence, respectively. No increase in CAT levels was observed in transfected cells which lack EBF (SP2 myeloma, BW5147 thymoma and NIH3T3 fibroblasts, latter not shown). As a positive control, separate transfections were performed with the plasmid MUFwt4, which contains four copies of the proximal *mb-1* promoter sequence (−59/−38; Travis *et al.*, 1991b). As expected, the *mb-1* promoter elements in MUFwt4 increased CAT levels in all four cell lines. Taken together, these data indicate that the ability of the distal *mb-1* promoter fragment to activate the c-*fos* promoter is dependent upon an intact EBF binding site and is detected only in cells that contain EBF.

We next examined whether an EBF binding site by itself is sufficient for conferring early B cell specificity upon a

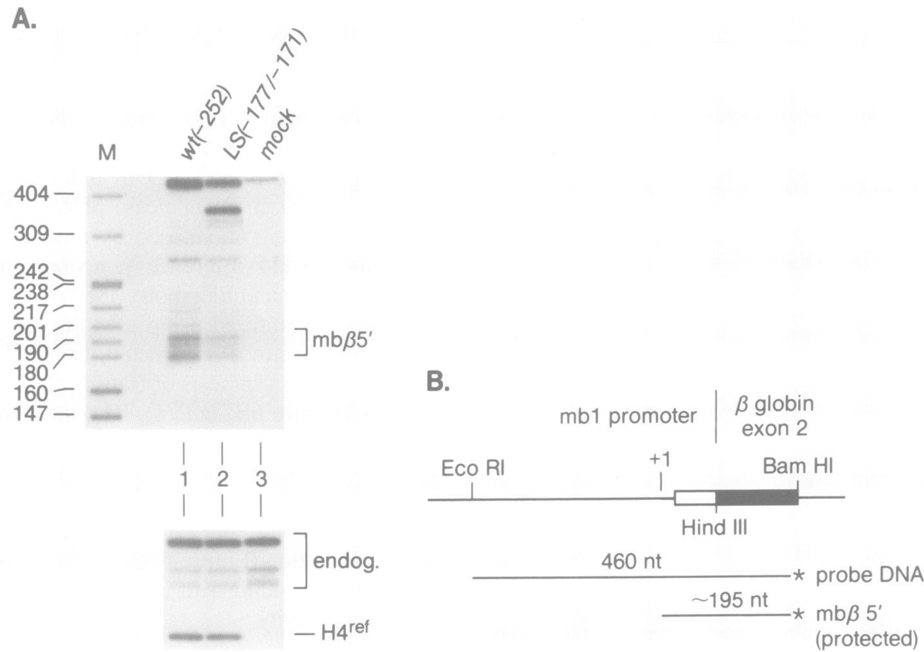

Fig. 6. The EBF binding site is required for maximal *mb-1* promoter function in transfected pre-B cells. Wild-type or mutated versions of the wt(−252)βPy plasmid, whch contains the intact *mb-1* promoter (−252/+70) linked to a β-globin reporter gene, a histone H4 reference gene for monitoring transfection efficiency, and the polyomavirus origin of replication were transiently transfected into the pre-B cell line PD36. Total cellular RNA was isolated at 48 h post-transfection and analyzed by an S1 nuclease protection assay. Mutations introduced in the *mb-1* distal promoter region in LS(−177/−171) are shown in Figure 4A. (**A**) S1 nuclease protection assay of RNA from transfected PD36 cells. Mock indicates an S1 nuclease protection assay performed with untransfected PD36 cell RNA. Upper panel: specifically initiated transcripts were detected with the *mb-1* promoter−β-globin probe (B) and are bracketed together as mbβ5′. The bracket also indicates the approximate window scanned by the PhosphorImager to estimate the number of specific transcripts. Strong bands which migrate more slowly than the correctly initiated transcripts represent protected read-through transcripts which initiate upstream of the *mb-1* promoter sequences. Lower panel: transcripts from the H4 reference gene ($H4^{ref}$) and from the endogenous H4 gene (endog.) were detected with an H4-specific probe (Grosschedl and Baltimore, 1985). (**B**) Structure of the *mb-1* promoter−β-globin probe. An *Eco*RI−*Hin*dIII fragment comprising the *mb-1* promoter (−252/+70, shown as *Eco*RI−*Hin*dIII) is fused to murine β-globin coding sequence (filled box) as in the plasmid wt(−252)βPy (Travis *et al.*, 1991b). The sizes of the 5′-end labeled probe and specifically protected fragments (approximate, due to multiple initiation sites) are indicated.

heterologous promoter. One or six copies of a 30 bp oligonucleotide (−183/−154) containing the wild-type or m4 mutated EBF binding site were inserted 5′ of the c-*fos* promoter in the fosCAT plasmid, or at a similar position in a β-globin promoter/reporter plasmid. Transfection of these plasmids into PD36 or M12 B lineage cell lines indicated that the EBF binding site by itself, even when multimerized, cannot increase transcription from either heterologous promoter (data not shown). As a control, a single copy of the proximal *mb-1* promoter region segment (−59/−38) increased c-*fos* promoter activity 10-fold in transfected PD36 cells (data not shown; Travis *et al.*, 1991b). Thus, the activation of transcription by the EBF binding site is dependent upon other sequences present in the distal *mb-1* promoter region.

Discussion

The identification of EBF represents an important step towards defining the mechanisms that specify the developmental expression pattern of the *mb-1* gene and possibly other B cell-specific genes. EBF is a B lineage-specific and temporally regulated nuclear factor that is expressed only in pre-B and surface IgM-positive B cells. The cell type-specific distribution of EBF coincides with the expression pattern of the endogenous *mb-1* gene. A functional role for EBF in *mb-1* gene expression is suggested by the observation that mutation of the EBF binding site decreases transcription from the *mb-1* promoter. Moreover, in the context of the distal *mb-1* promoter region, the EBF binding site can confer upon a heterologous promoter a cell type-specific pattern of activity that is characteristic of the *mb-1* promoter. In addition, EBF specifically recognizes a sequence within the MHC class II A_{α}^{d} promoter that is identical to the core of the *mb-1* promoter EBF binding site. The A_{α}^{d} EBF binding site is present between the regulatory X and Y boxes, suggesting that EBF may also have a role in the control of MHC class II gene expression. Taken together, these data establish EBF as a transcriptional regulator of early B cell-specific gene expression.

EBF is distinct from other known DNA binding proteins that regulate gene expression in B cells. The developmental regulation and cell type distribution of EBF is paralleled by the B cell specific activator protein, or BSAP (Barberis *et al.*, 1990). Because the expression pattern of the *mb-1* gene matches that of BSAP, the *mb-1* gene is a potential target for regulation by this factor. Therefore, it was important to define whether these two factors represent related DNA binding activities. Three lines of evidence suggest that EBF is distinct from BSAP. First, the core sequence of the EBF binding site does not resemble the sea urchin late histone promoter sequences recognized by BSAP. Moreover, cross-competition of these nucleotide sequences for protein binding was not detected *in vitro*. Second, when probes of identical

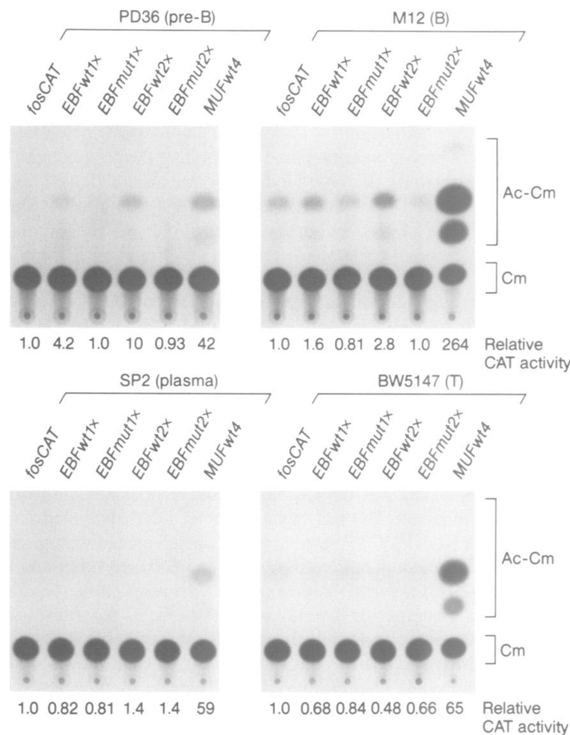

Fig. 7. EBF-dependent activation of the c-*fos* promoter by the distal *mb-1* promoter region in early stage B cell lines. The fosCAT plasmid or derivative plasmids with one or two copies of the wild-type (EBFwt1X and EBFwt2X) or mutated (EBFmut1X and EBFmut2X) *mb-1* promoter fragments (−217/−113) inserted at −56 of the c-*fos* promoter were transfected together with the RSV−luciferase plasmid as an internal standard. The mutated *mb-1* distal promoter region fragment contains the LS(−177/−171) mutation (Figure 4A). Extracts for CAT determinations were normalized to luciferase activity values prior to the assay of CAT activity. A representative CAT assay is shown for each cell line. Acetylated chloramphenicol (Ac-Cm) and unreacted chloramphenicol (Cm) were separated by TLC and the relative CAT levels were estimated as (Ac-Cm c.p.m.)/(Ac-Cm + Cm c.p.m.) and normalized to the CAT activity measured for the fosCAT plasmid (1.0). The variability between independent experiments was <20%.

length were used in an electrophoretic mobility shift assay, the rate of migration of the BSAP complex was much faster than that of the EBF complex. Finally, UV light covalently cross-linked a single 45−50 kDa polypeptide to the BSAP binding site (Barberis *et al.*, 1990; this study), whereas, two polypeptides of 70−75 kDa and 80−85 kDa were cross-linked to the EBF recognition site. The identical developmental expression patterns of these two nuclear factors suggests that EBF and BSAP may coordinately regulate a common set of genes. Alternatively, EBF or BSAP may control the expression of the other factor as part of a regressive hierarchy of transcriptional regulation.

The initial characterization of EBF suggests several notable features. First, the factor appears to be comprised of at least two polypeptides. Based on the high efficiency of cross-linking to the DNA probe, we anticipate that both polypeptides make tight contacts with DNA. Second, the binding of EBF to DNA includes contacts with both the major and minor grooves and is inhibited by methylation at nine contiguous nucleotides. Third, EBF regulates *mb-1* transcription in a context-dependent manner. EBF can bind its recognition site within a 30 bp oligonucleotide *in vitro*, yet it requires additional sequences from the *mb-1* promoter to enhance transcription from the c-*fos* promoter. These data suggest that the activation of transcription by EBF is dependent on other factors that bind distal promoter region sequences. In one possible scheme, neither factor alone may be able to enhance transcription, but the combination of the two is sufficient for activation. As an example, the regulation of basal transcription of the yeast *HIS4* gene requires both the BAS1 and BAS2 gene products, which appear to bind DNA non-cooperatively (Tice-Baldwin *et al.*, 1989). Alternatively, EBF may stimulate transcription indirectly by augmenting the binding and/or function of a transcriptional activator. The context-dependence of transcriptional enhancement by EBF is reminiscent of that of the pre-B and T cell-specific factor LEF-1/TCF-1α (Waterman and Jones, 1990; Waterman *et al.*, 1991; Travis *et al.*, 1991a). Multimerized binding sites for LEF-1 do not enhance transcription, but a single site can augment enhancer function in association with other factor binding sites within the T cell receptor α enhancer. Therefore, the function of the EBF and LEF-1 binding sites contrasts with that of binding sites for other lymphoid-specific factors, including NFκB (Wirth and Baltimore, 1988), PU.1 (Klemsz *et al.*, 1990) and BSAP (Barberis *et al.*, 1990), which were shown to increase transcription from heterologous promoters independently of sequences that normally flank the cognate binding sites.

Our experiments indicate that the EBF binding site in the intact *mb-1* promoter is important for early B cell-specific transcription, because point mutations in the EBF binding site reduce promoter activity 3-fold. In contrast, we have observed that the deletion of distal promoter sequences that include the EBF binding site (−252/−159) did not affect the level of transcription relative to the intact promoter (Travis *et al.*, 1991b). It is possible that the deletion of such a large part of the *mb-1* promoter region has multiple consequences, including the removal of negatively acting sequences which obscure the effect of deleting the EBF binding site. The minimal *mb-1* promoter fragment lacking the entire distal promoter region retains the cell type specificity of the intact promoter (Travis *et al.*, 1991b), suggesting regulation by additional, as yet unidentified, cell type-specific factors. Therefore, the *mb-1* promoter appears to contain multiple sequences that specifically regulate transcriptional activity in early B cells. Within the proximal *mb-1* promoter region, we have identified two factor binding sites (located between −59 and −38) that are important for *mb-1* promoter activity. Mutation at either of these sites decreases transcription 4-fold, and deletion of both sites reduces promoter function at least 20-fold. However, when inserted upstream of a heterologous promoter, the two proximal promoter region binding sites can enhance transcription in all cell types, and therefore these sites alone are unlikely candidates for cell type-specific control elements. In agreement with these observations, two factors (Sp1 and MUF2) which bind this region *in vitro* are present in cells that do not express *mb-1* (Travis *et al.*, 1991b). Thus, although EBF regulates the level of *mb-1* transcription in early stage B cells, the ultimate cell type specificity of the promoter may require additional factors that have not yet been identified.

In conclusion, we have identified EBF as a cell type-specific regulator of gene expression in the early stages of B cell differentiation. In addition to its role for the regulation

of *mb-1* gene expression, we suspect that EBF participates in the regulation of other genes involved in specifying the phenotype of early stage B cells.

Materials and methods

Cell culture and nuclear extracts

S194/5.XXO.Bu1 (S194; American Type Culture Collection) cells were cultured in DMEM supplemented with 10% horse serum. WEHI-3 cells were grown in Joklik's medium supplemented with 10% fetal bovine serum (FBS). The NIH3T3 fibroblast line was cultured in DMEM supplemented with 10% FBS. All other lymphoid cell lines were cultured in RPMI 1640 supplemented with 10% FBS and 50 μM 2-mercaptoethanol.

Nuclear extracts were prepared by the method of Dignam *et al.* (1983) with the addition of the following protease inhibitors: 1 μg/ml pepstatin, 5 μg/ml leupeptin, 1 mM ϵ-amino-caproic acid, 10 μM *n*-tosyl-L-phenylalanine-chloromethyl-ketone, 1 μM aprotinin and 5 μg/ml chymostatin. Dimethyl sulfoxide-treated (50 h) MEL nuclear extract was kindly provided by L.Stuve. Splenic B cell nuclear proteins were extracted from an enriched population of B cells isolated from the spleens of 6–8 week old C57Bl6 mice as described (Gold *et al.*, 1990).

Plasmid construction and in vitro mutagenesis

Plasmids pwt(−252)βPy, pBluMβ and pM9-4S1 are described in Travis *et al.* (1991b). The *mb-1* promoter plasmid pmb1BH was made by subcloning the 322 bp *Bam*HI–*Hin*dIII fragment of pM9-4S1 into pUC19 digested with *Bam*HI and *Hin*dIII.

In vitro mutagenesis of the distal region of the *mb-1* promoter in pBluMβ was performed using the method of Kunkel (1985) and the primer 5′-CTGGCCACAATTCCCGCTGCAGTCTCTCTAGTGCGTG-3′ to make BluLS(−177/−171). The 322 bp *Bam*HI–*Hin*dIII fragment of BluLS(−177/−171) was excised and ligated in place of the wild-type sequence of pwt(−252)βPy to make LS(−177/−171)Mβ. The 322 bp *Bam*HI–*Hin*dIII fragment of BluLS(−177/−171) was subcloned into pUC19 digested with *Bam*HI and *Hin*dIII to make pmb1LS(−177/−171).

The fosCAT test plasmid (p301-56) was kindly provided by M.Gilman (Gilman *et al.*, 1986). To generate the *mb-1* promoter wild-type (EBFwt1X and EBFwt2X) and mutant (EBFmut1X and EBFmut2X) plasmids, the 105 bp *Stu*I–*Nde*I fragments of pmb1BH and pmb1LS(−177/−171) were isolated following growth of these plasmids in the dam$^-$ dcm$^-$ *Escherichia coli* strain GM48 (Marinus, 1973). The *Stu*I–*Nde*I fragments were blunt-ended using Klenow DNA polymerase, and one or two copies were subcloned into the *Sal*I site (similarly blunted) at −56 of the c-*fos* promoter in fosCAT.

Synthetic oligonucleotides for insertion in the fosCAT vector were annealed, phosphorylated with T4 polynucleotide kinase and cold ATP, and ligated directly into the blunted fosCAT vector, or ligated for multimerization before inactivating ligase and digesting with *Xba*I and *Spe*I to isolate head to tail concatamers. Concatamers were then cloned into the *Xba*I site of Bluescript (Stratagene) for sequencing, cut out with *Xba*I and *Spe*I, blunt-ended with Klenow DNA polymerase, and insertion into the blunted *Sal*I site of fosCAT. Sequences of wild-type and m4 mutated oligonucleotide pairs were:

5′-CTAGAGAGAGAGACTCAAGGGAATTGTGGCCAGCCA-3′
3′ TCTCTCTCTGAGTTCCCTTAACACCGGTCGGTGATC-5′

5′-CTAGAGAGAGAGACTCAACCGTTTTGTGGCCAGCCA-3′
TCTCTCTCTGAGTTGGCAAAACACCGGTCGGGTGAT-5′.

DNase I footprint assay

The *mb-1* promoter plasmid pmb1BH was digested with either *Bam*HI or *Nde*I, dephosphorylated with calf intestinal phosphate (CIP), 5′-end labeled with [γ-^{32}P]ATP (Amersham) using T4 polynucleotide kinase, and then cleaved with either *Nde*I or *Bam*HI to label one end of the non-coding or coding strand, respectively. Probes were purified on native polyacrylamide gels to obtain the (−252/−113) probe fragments. Limited digestion reactions with DNase I (DPRF-grade; Worthington Biochemicals) were performed as previously reported (Travis *et al.*, 1991a) with the following modifications. Binding reactions (30 μl) were carried out at 20°C for 30 min with 2 fmol of labeled probe in the presence or absence of 50 μg nuclear protein, and contained 300 μg/ml poly(dI-dC)–poly(dI-dC) (Pharmacia) as a non-specific competitor. Digestion reactions were performed at 20°C with 10–20 μg/ml (without nuclear extract) or 80–120 μg/ml (with nuclear protein) DNase I for 90 s. Reactions were stopped and processed as previously reported. G and G+A sequencing ladders of the probes were prepared as described (Maniatis *et al.*, 1982).

Electrophoretic mobility shift and methylation interference assays

Plasmid fragment probes were prepared by digesting the plasmid pmb1BH with appropriate enzymes, de-phosphorylating 5′ ends with CIP, and labeling with [γ-^{32}P]ATP and T4 polynucleotide kinase, and filling in recessed 3′ ends with Klenow DNA polymerase and cold dNTPs. Synthetic oligonucleotide probes were similarly labeled, omitting the CIP treatment. 0.5–1 fmol of labeled probe was incubated together with nuclear extract and competitor oligonucleotides in a final volume of 20 μl of a buffer consisting of 10 mM HEPES (pH 7.9), 70 mM KCl, 1 mM DTT, 4% glycerol, 1 mM EDTA, 2.5 mM $MgCl_2$ for 30 min at 20°C. Poly(dI-dC)–poly(dI-dC) was added at 200 μg/ml as a non-specific competitor. 10 μl of each reaction was electrophoresed through a native 6% polyacrylamide gel in 25 mM Tris, 190 mM glycine and 1 mM EDTA at 4°C. Gels were dried and subjected to autoradiography.

For methylation interference analysis, the 118 bp *Dde*I fragment of pmb1BH(−243/−126) was labeled on the coding or non-coding strands by filling in recessed 3′ ends with Klenow DNA polymerase and [α-^{32}P]dCTP or [α-^{32}P]dGTP respectively. Labeled probes were partially methylated with dimethyl sulfate (Siebenlist and Gilbert, 1980) and used in a scaled up (20×) electrophoretic mobility shift assay. After separation of protein–DNA complexes (bound) and free probe as described above, probe DNA was eluted from gel slices in 10 mM Tris (pH 7.5), 0.5 M NaOAc, 10 mM $MgCl_2$, 0.1% SDS for 3 h at 37°C, extracted with an equal volume of phenol–chloroform (1:1) and chloroform, precipitated with ethanol and rinsed with 70% ethanol. Samples containing the partially methylated probe were then cleaved in 100 μl 1 M piperidine at 95°C for 30 min, then lyophilized, redissolved in 100 μl water and again lyophilized three times to dryness. Interference with binding by methylated adenines can be detected in this assay, but much less efficiently than interference by methylated G residues (Baldwin, 1988). Equal amounts of radioactivity were subjected to electrophoresis on a 6% polyacrylamide–urea gel and visualized by autoradiography.

UV cross-linking assay

The UV cross-linking assay was essentially performed according to the method of Ballard *et al.* (1989). To make the substituted DNA probes, synthetic oligonucleotides were annealed and extended using Klenow DNA polymerase, [α-^{32}P]dATP, [α-^{32}P]dCTP, 5-Br-deoxyUTP (Boehringer-Mannheim) and unlabeled dGTP. The *mb-1* promoter (EBF binding site) probe was prepared by annealing and extending the oligonucleotides 5′-GAGAGAGACTCAAGGGAATTGTGGCCAGCC-3′ and 5′-GGCTGGCCA-3′; the sea urchin late histone H2A-2.2 (BSAP binding site) probe was similarly prepared using 5′-GGTTGTGACGCAGCGGTGGGTGACGACTGT-3′ and 5′-ACAGTCGTC-3′. DNA binding reactions (20 μl) were initiated by mixing 25–35 fmol of labeled probe with 5 μg of PD36 nuclear extract under the conditions described for other electrophoretic mobility shift assays. Following native gel electrophoresis, factor–DNA complexes were irradiated *in situ* at 4°C for various times using a 300 nM UV transilluminator (UVP, Inc.). Factor–DNA adducts were isolated and electrophoresed as described (Ballard *et al.*, 1989).

Transfection, S1 nuclease analysis, CAT and luciferase assays and quantification

Short-term DNA transfections into cell lines were performed as described in Grosschedl and Baltimore (1986) with the modifications of Travis *et al.* (1991a). Isolation of mRNA, probe preparation and the S1 nuclease assays were performed as described (Travis *et al.*, 1991b). CAT and transfection control luciferase assays are described in Travis *et al.* (1991a).

Following thin layer chromatography, the quantification of [^{14}C]chloramphenicol acetylation was made by direct measurement on a PhosphorImager system (Molecular Dynamics).

Acknowledgements

The authors appreciate the contributions of Debbie Law for assisting with spleen cell isolation and FACS analysis, Marie Doherty for synthesizing oligonucleotides, Peter Engler for providing bacterial strains and Adam Amsterdam, Steve Hardy and Laura Stuve for sharing nuclear extracts. The authors would also like to thank Bill Forrester, Tony DeFranco and Keith Yamamoto for their critical review of this manuscript. J.Hagman was

supported by a postdoctoral training grant from the National Institute of Health and by a postdoctoral fellowship awarded by the Leukemia Society of America. This work was supported by funds from the Howard Hughes Medical Institute.

References

Baldwin,A.S.Jr (1988) In *Current Protocols in Molecular Biology*, Vol. 2, 12.3.5.

Ballard,D.W., Bohnlein,E., Hoffman,J.A., Bogerd,H.P., Dixon,E.P., Franza,B.R. and Greene,W.C. (1989) *New Biol.*, **1**, 83–92.

Barberis,A., Widenhorn,K., Vitelli,L. and Busslinger,M. (1990) *Genes Dev.*, **4**, 849–859.

Benoist,C. and Mathis,D. (1990) *Annu. Rev. Immunol.*, **8**, 681–715.

Bergman,Y., Rice,D., Grosschedl,R. and Baltimore,D. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 7041–7045.

Cambier,J.C. (1991) *Curr. Biol.*, **1**, 25–27.

Campbell,K.S. and Cambier,J.C. (1990) *EMBO J.*, **9**, 441–448.

Celada,A., Shiga,M., Imagawa,M., Kop,J. and Maki,J.A. (1988) *J. Immunol.*, **140**, 3995–4002.

Dedrick,R.L. and Jones,P.P. (1990) *Mol. Cell. Biol.*, **10**, 593–604.

Dignam,J.D., Lebovitz,R.M. and Roeder,R.G. (1983) *Nucleic Acids Res.*, **11**, 1475–1487.

Gilman,M.Z., Wilson,R.N. and Weinberg,R.A. (1986) *Mol. Cell. Biol.*, **6**, 4305–4316.

Gold,M.R., Law,D.A. and DeFranco,A.L. (1990) *Nature*, **345**, 810–813.

Gravallese,E.M., Boothby,M., Smas,C.M. and Glimcher,L. (1989) *Mol. Cell. Biol.*, **9**, 3184–3192.

Grosschedl,R. and Baltimore,D. (1985) *Cell*, **41**, 885–897.

Hombach,J., Leclercq,L., Radbruch,A., Rajewsky,K. and Reth,M. (1988) *EMBO J.*, **7**, 3451–3456.

Hombach,J., Tsubata,T., Leclercq,L., Stappert,H. and Reth,M. (1990) *Nature*, **343**, 760–762.

Klemsz,M.J., McKercher,S.R., Celada,A., Van Beveren,C. and Maki,R.A. (1990) *Cell*, **61**, 113–124.

Kudo,A. and Melchers,M. (1987) *EMBO J.*, **6**, 2267–2272.

Kunkel,T.A. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 488–492.

Laemmli,U.K. (1970) *Nature*, **277**, 680–685.

Landau,N.R., St. John,T.P., Weissman,I.L., Wolf,S.C., Silverstone,A.E. and Baltimore,D. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 5836–5840.

Lenardo,M.J. and Baltimore,D. (1989) *Cell*, **58**, 227–229.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp. 475–476.

Marinus,M.G. (1973) *Mol. Gen. Genet.*, **127**, 47–55.

Max,E.E., Maizel,J.V.Jr and Leder,P. (1981) *J. Biol. Chem.*, **256**, 5116–5120.

Reth,M. (1989) *Nature*, **338**, 383–384.

Sakaguchi,N. and Melchers,F. (1986) *Nature*, **324**, 579–582.

Sakaguchi,N., Berger,C.N. and Melchers,F. (1986) *EMBO J.*, **5**, 2139–2147.

Sakaguchi,N., Kashiwamura,S.-I., Kimoto,M., Thalmann,P. and Melchers,F. (1988) *EMBO J.*, **7**, 3457–3464.

Siebenlist,U. and Gilbert,W. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 122–126.

Singh,H.R., Sen,R., Baltimore,D. and Sharp,P.A. (1986) *Nature*, **319**, 154–158.

Staudt,L.M. and Lenardo,M.J. (1991) *Annu. Rev. Immunol.*, **9**, 373–398.

Staudt,L.M., Clerc,R.G., Singh,H., LeBowitz,J., Sharp,P.A. and Baltimore,D. (1988) *Science*, **241**, 577–580.

Tice-Baldwin,K., Fink,G.R. and Arndt,K.T. (1989) *Science*, **246**, 931–935.

Travis,A., Amsterdam,A., Belanger,C. and Grosschedl,R. (1991a) *Genes Dev.*, **5**, 880–894.

Travis,A., Hagman,J. and Grosschedl,R. (1991b) *Mol. Cell. Biol.*, **11**, in press.

Waterman,M.L. and Jones,K.A. (1990) *New Biol.*, **2**, 621–636.

Waterman,M.L., Fisher,W.H. and Jones,K.A. (1991) *Genes Dev.*, **5**, 656–669.

Wirth,T. and Baltimore,D. (1988) *EMBO J.*, **7**, 3109–3113.

Received on June 28, 1991