The EMBO Journal vol.15 no.8 pp.1971–1982, 1996

C-terminal activating and inhibitory domains determine the transactivation potential of BSAP (Pax-5), Pax-2 and Pax-8

Petra Dörfler and Meinrad Busslinger[1]

Research Institute of Molecular Pathology, Dr Bohr-Gasse 7, A-1030 Vienna, Austria

[1]Corresponding author

***Pax-5* encodes the transcription factor BSAP which plays an essential role in early B cell development and midbrain patterning. In this study we have analysed the structural requirements for transcriptional activation by BSAP. *In vitro* mutagenesis and transient transfection experiments indicate that the C-terminal serine/threonine/proline-rich region of BSAP contains a potent transactivation domain of 55 amino acids which is active from promoter and enhancer positions. This transactivation domain was found to be inactivated by a naturally occurring frameshift mutation in one *PAX-5* allele of the acute lymphoblastic leukemia cell line REH. The function of the transactivation domain is negatively regulated by adjacent sequences from the extreme C-terminus. The activating and inhibitory domains function together as an independent regulatory module in different cell types as shown by fusion to the GAL4 DNA binding domain. The same arrangement of positively and negatively acting sequences has been conserved in the mammalian Pax-2 and Pax-8, the zebrafish Pax-b as well as the sea urchin Pax-258 proteins. These data demonstrate that the transcriptional competence of a subfamily of Pax proteins is determined by a C-terminal regulatory module composed of activating and inhibitory sequences.**

Keywords: BSAP (Pax-5)/negative regulation/Pax-2/Pax-8/transactivation

Introduction

The paired box containing (*Pax*) genes constitute a small family of conserved developmental control genes which can be grouped into different subclasses according to structural similarities (reviewed by Noll, 1993). The nine mammalian *Pax* genes are known to encode transcription factors with important functions in embryonic pattern formation, which is best illustrated by their association with mouse developmental mutants and human disease syndromes (reviewed by Stuart *et al.*, 1994). We have previously demonstrated that the B cell-specific transcription factor BSAP is encoded by the *Pax-5* gene (reviewed by Busslinger and Urbánek, 1995). BSAP is expressed at all stages of B cell development except in terminally differentiated plasma cells (Barberis *et al.*, 1990) and regulates the *CD19* gene which codes for a co-stimulatory molecule of the B cell receptor (Kozmik *et al.*, 1992). In addition to all B-lymphoid tissues, the *Pax-5* gene is also expressed in the developing midbrain and adult testis of the mouse (Adams *et al.*, 1992). In accordance with this expression pattern, gene inactivation in the mouse germline revealed that *Pax-5* plays a important role in B-lymphopoiesis and midbrain development (Urbánek *et al.*, 1994).

The transcription factors of the Pax family recognize their target genes via the DNA binding function of the paired domain (Chalepakis *et al.*, 1991; Treisman *et al.*, 1991). Detailed mutational analysis of BSAP (Pax-5) revealed a bipartite structure of the paired domain and its binding sites, and led to the identification of a consensus recognition sequence bound by all Pax proteins analysed to date (Czerny *et al.*, 1993). Initially, Pax proteins were assumed to be transcription factors solely based on their nuclear localization (Bopp *et al.*, 1989) and DNA binding potential. Subsequently, cell transfection experiments demonstrated transcriptional activity for all members of the Pax protein family analysed: Pax-1 (Chalepakis *et al.*, 1991), Pax-2 (Fickenscher *et al.*, 1993), Pax-3 (Chalepakis *et al.*, 1994), BSAP (Pax-5) (Barberis *et al.*, 1990), Pax-6 (Glaser *et al.*, 1994; Czerny and Busslinger, 1995), Pax-7 (Schäfer *et al.*, 1994) and Pax-8 (Zannini *et al.*, 1992; Kozmik *et al.*, 1993). Most Pax proteins end in distinct serine/threonine/proline-rich sequences (PST-rich) which are thought to constitute transactivation domains, since their elimination by gross deletion (Fickenscher *et al.*, 1993; Chalepakis *et al.*, 1994; Czerny and Busslinger, 1995) or replacement by alternative splicing (Kozmik *et al.*, 1993) reduces or abrogates the transactivation function of Pax proteins. However, none of these transactivation domains has been defined and characterized in detail by *in vitro* mutagenesis, yet detailed knowledge of the transactivation function will be required to further our understanding of how transcription factors of the Pax family regulate gene expression during development and differentiation.

Here we present a detailed structure–function analysis of BSAP (Pax-5) which has led to the identification of a potent transactivation domain of 55 amino acids located within the C-terminal PST-rich sequences. This domain exerts its activating function from a promoter as well as an enhancer position and is subject to strong negative regulation by adjacent sequences from the extreme C-terminus. Interestingly, a naturally occurring frameshift mutation which inactivates the transactivation domain of BSAP was identified in one *PAX-5* allele of a B-lymphoid cell line derived from a patent with acute lymphoblastic leukemia (ALL). Comparative studies furthermore revealed that both the transactivating sequences and the adjacent inhibitory domain are structurally and functionally conserved in all members of the Pax protein subfamily consisting of the mammalian Pax-2, Pax-5, Pax-8, the zebrafish Pax-b and the distantly related sea urchin Pax-258

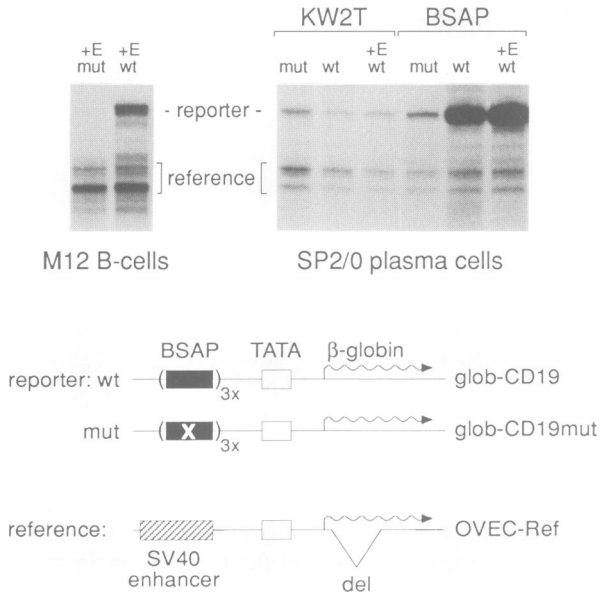

Fig. 1. Activation of reporter gene transcription by endogenous and exogenous BSAP. Three copies of the high-affinity BSAP binding site CD19-2(A-ins) (Czerny *et al.*, 1993) or a mutant derivative differing by two nucleotide substitutions, were inserted upstream of the TATA box of the β-globin gene in OVEC-S to generate the reporter constructs glob–CD19 and glob–CD19mut, respectively (see Materials and methods). Where indicated, these reporter constructs contained the SV40 enhancer (+E) downstream of the β-globin gene. The wild-type (wt) or mutant (mut) reporter constructs (9 μg) were transiently transfected into cells of the murine B cell line M12.4.1 together with the reference gene OVEC-Ref (2 μg) (Westin *et al.*, 1987). Expression of the β-globin reporter genes was monitored by RNase protection assay. Cells of the murine plasma cell line SP2/0 were additionally transfected with a BSAP expression plasmid or the empty expression vector pKW2T (2.5 μg each).

proteins. These data indicate that the transcriptional activity of a subset of Pax proteins is determined by a C-terminal regulatory module consisting of activating and inhibitory sequences.

Results

Development of a reporter system to study the transactivation function of BSAP

The promoters of the BSAP-regulated gene *CD19* and the putative target genes *blk*, V_{preB} and *Iε* are weakly active in transiently transfected B cells and can only be minimally stimulated by ectopic BSAP expression in heterologous cell types (reviewed by Busslinger and Urbánek, 1995). Common to all of these genes is the fact that their promoters do not contain a TATA box and consequently initiate transcription at heterogeneous start sites. These promoters could therefore not be used to generate efficient reporter constructs for investigating the transactivation function of BSAP in cell transfection experiments. Hence, we created an artificial promoter by inserting three copies of a high-affinity BSAP binding site derived from the CD19 promoter [CD19-2(A-ins)] upstream of the TATA box of the rabbit β-globin gene (Figure 1; wt reporter glob–CD19). This reporter construct was transiently transfected into several B cell lines together with the internal reference gene OVEC-Ref containing a deletion in the β-globin leader sequences (Westin *et al.*, 1987), and β-globin transcripts were detected by RNase protection analysis. As illustrated in Figure 1, this reporter gene was efficiently transcribed in the murine B cell line M12.4.1 most probably due to activation by endogenous BSAP. To directly demonstrate BSAP-dependent regulation, we mutated the reporter gene by introducing two point mutations into each BSAP recognition sequence (reporter glob–CD19mut, Figure 1) which interfered with protein binding in electrophoretic mobility shift assays (P.Dörfler, unpublished data). These mutations indeed prevented activation of reporter gene transcription by BSAP (Figure 1). Analogous results were obtained in BSAP-negative plasma cell lines by co-transfection of the two reporter plasmids with a BSAP expression vector. As shown for the murine plasma cell line SP2/0 (Figure 1), transcription of the wt reporter gene was strongly and specifically activated by BSAP compared with expression of the mutant reporter gene or with the basal-level expression observed in the presence of the empty expression vector pKW2T. In B cells, strong transactivation by endogenous BSAP was reproducibly seen only in the presence of a heterologous enhancer (SV40 enhancer) inserted downstream of the β-globin reporter gene. Such an enhancer was not necessary for activation in the plasma cell line SP2/0 (lanes +E in Figure 1) and thus all further analyses in plasma cells were performed with the wt reporter gene glob–CD19 lacking the SV40 enhancer. Moreover, analysis of five different plasma cell lines demonstrated that the transactivation potential of BSAP was highest in SP2/0 cells (P.Dörfler, unpublished data). Hence, SP2/0 cells were initially used to define the transactivation function of BSAP.

The C-terminal sequences of BSAP harbour a potent transactivation domain which is active from promoter and enhancer positions

To elucidate the BSAP domain(s) required for transactivation, we generated a series of C-terminal and internal deletion mutants (Figure 2, constructs B2–B18) which we analysed by transient transfection in SP2/0 cells, utilizing the above described reporter system (Figure 3). Deletion of up to 33 amino acids from the C-terminus of BSAP (constructs B2–B4) did not reduce the transactivation potential compared with the full-length protein (B1; Figure 3A). However, deletion of additional 13 amino acids in mutant B5 drastically reduced BSAP-dependent transcription (Figure 3A). The same reduction was also observed for all further C-terminal deletion mutants (B6–B9; Figure 3A). As this result could be explained by failed synthesis or instability of the truncated BSAP peptides, we analysed nuclear extracts prepared from the same transfected cells by electrophoretic mobility shift assay (EMSA). As shown in Figure 3A, all BSAP polypeptides were expressed at comparable levels, thus indicating that the observed differences in transactivation did not result from abnormal expression of the mutant proteins. In addition, immunocytochemistry with anti-paired domain antibodies demonstrated that transcriptionally inactive deletion mutants such as B9 were localized in the nucleus like the full-length protein (B1; Figure 3C). The immunofluorescence analysis was performed with transiently transfected murine COP-8 fibroblasts (Tyndall *et al.*, 1981) which allow a better distinction between nuclear and cytoplasmic compartments

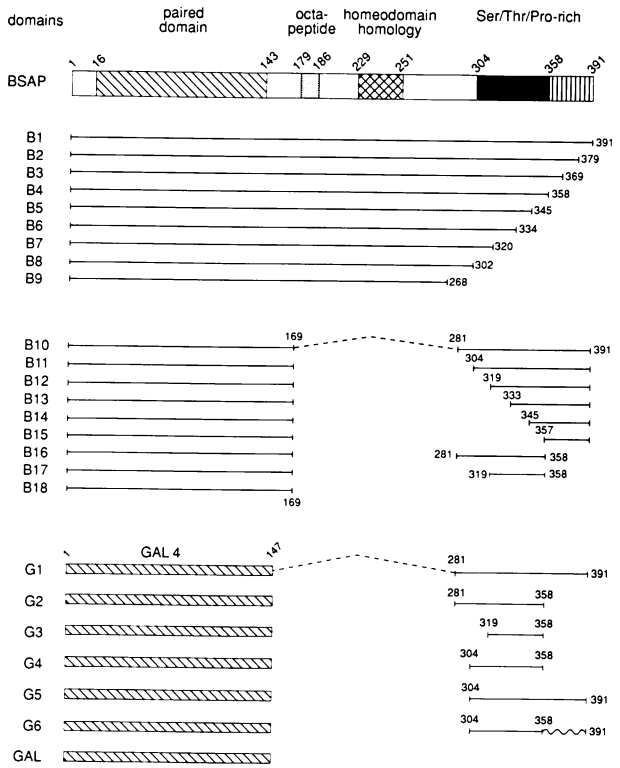

Fig. 2. Schematic diagram of BSAP deletion mutants and chimeric GAL4–BSAP proteins. The structural organization of human BSAP (B1) is shown together with the extent of the various deletions engineered into the mutant BSAP proteins B2–B18. Numbers refer to the corresponding amino acid positions of human BSAP (Adams *et al.*, 1992). The DNA binding domain of the yeast GAL4 protein (amino acids 1–147; indicated by a hatched bar) was fused to C-terminal sequences of BSAP in the chimeric proteins G1–G6. A wavy line indicates a shift in reading frame near the C-terminus of protein G6 (for sequence see Figure 9). For generation of the different constructs see Materials and methods.

than plasma cells. These control experiments thus indicated that the reduction in reporter gene transcription was indeed due to loss of the transactivation function in the BSAP deletion mutants. The C-terminal deletion analysis therefore placed the C-terminal boundary of the BSAP transactivation domain near amino acid 358 (construct B4).

In order to delimit the N-terminal boundary of the BSAP transactivation domain, we generated an internal deletion by fusing the PST-rich C-terminal sequences at codon 281 to the paired domain which on its own lacks any transactivation function (construct B18). Although this large internal deletion removed the conserved octapeptide motif, the homeodomain homology region and additional serine/threonine-rich sequences (see Figure 2), the respective mutant protein B10 exhibited similar transcriptional activity as full-length BSAP (Figure 3B). Successive deletion of amino acids 281–319 in constructs B11 and B12 did not affect the transactivation function of BSAP in SP2/0 cells. In contrast, further deletion of the PST-rich sequences gradually decreased the transcriptional activity of the mutant proteins, whilst leaving their expression level unaltered (constructs B13–15; Figure 2B). These data therefore map the N-terminal boundary of the transactivation domain close to amino acid 319 (construct B12). The combined results of the N- and C-terminal deletion analyses predict that the sequences between amino acids 319 and 358 constitute the minimal transactivation domain of BSAP. Indeed, this 40 amino acid sequence linked to the paired domain (construct B17) activated reporter gene transcription as efficiently as full-length BSAP (Figure 3B) in SP2/0 cells.

We next investigated whether transactivation from a promoter or enhancer position would depend on the same or different BSAP sequences. A thymidine kinase (TK)–luciferase gene containing three BSAP binding sites inserted 2 kb downstream of the transcription start site of the TK promoter (Czerny and Busslinger, 1995) was co-transfected into SP2/0 cells together with expression plasmids directing the synthesis of full-length or mutant BSAP proteins. Interestingly, the minimal transactivation domain present in mutant B17 was already sufficient to activate luciferase expression as effectively as full-length BSAP (B1) (Figure 4). In both cases, transcription of the luciferase gene was increased ~4-fold relative to the basal activity of the TK promoter measured in the presence of the empty expression vector pKW2T. We conclude therefore that a minimal transactivation domain of 40 amino acids located in the C-terminal region of BSAP is sufficient in SP2/0 cells to mediate transcriptional activation both from promoter and enhancer positions.

The cell line REH derived from an ALL patient contains a frameshift mutation in one PAX-5 allele which eliminates the transactivation function of BSAP

Screening of established B-lymphoid cell lines for BSAP expression revealed a truncated form of this transcription factor in the cell line REH which was originally established from a patient with acute lymphoblastic leukemia (ALL) (Rosenfeld *et al.*, 1977). Nuclear extracts prepared from REH and control BJA-B cells were subjected to EMSA analysis with a paired domain binding probe. As shown in Figure 5A, REH cells synthesize both full-length BSAP and a truncated polypeptide in similar amounts, suggesting that one *PAX-5* allele may be mutated in these cells. BSAP cDNA was therefore amplified by PCR from REH cells, cloned into the expression vector pKW10 (Adams *et al.*, 1992) and analysed for protein expression in transfected COP-8 fibroblasts. Several cDNAs were isolated that encoded either full-length or truncated BSAP proteins, as exemplified by clones 1 and 2 (Figure 5A). DNA sequencing of clone 1 revealed that an additional C residue was inserted into a string of seven cytosines at positions 1032–1038 of the BSAP cDNA (Adams *et al.*, 1992). This single nucleotide insertion shifts the reading frame immediately upstream of the minimal transactivation domain, thus resulting in premature termination of translation after 18 additional amino acids (Figure 5C). The truncated protein of clone 1 consequently failed to activate BSAP-dependent reporter gene transcription, although it was as efficiently synthesized in transfected SP2/0 cells as the wild-type protein of clone 2 (Figure 5B). Interestingly, transcripts of the BSAP target gene *CD19* are markedly reduced in abundance in REH cells compared with most other human B cell lines (Kozmik *et al.*, 1992), which may reflect either the lower level of transactivation-competent BSAP and/or negative interference by the mutant protein in REH cells. This finding therefore suggests that the minimal

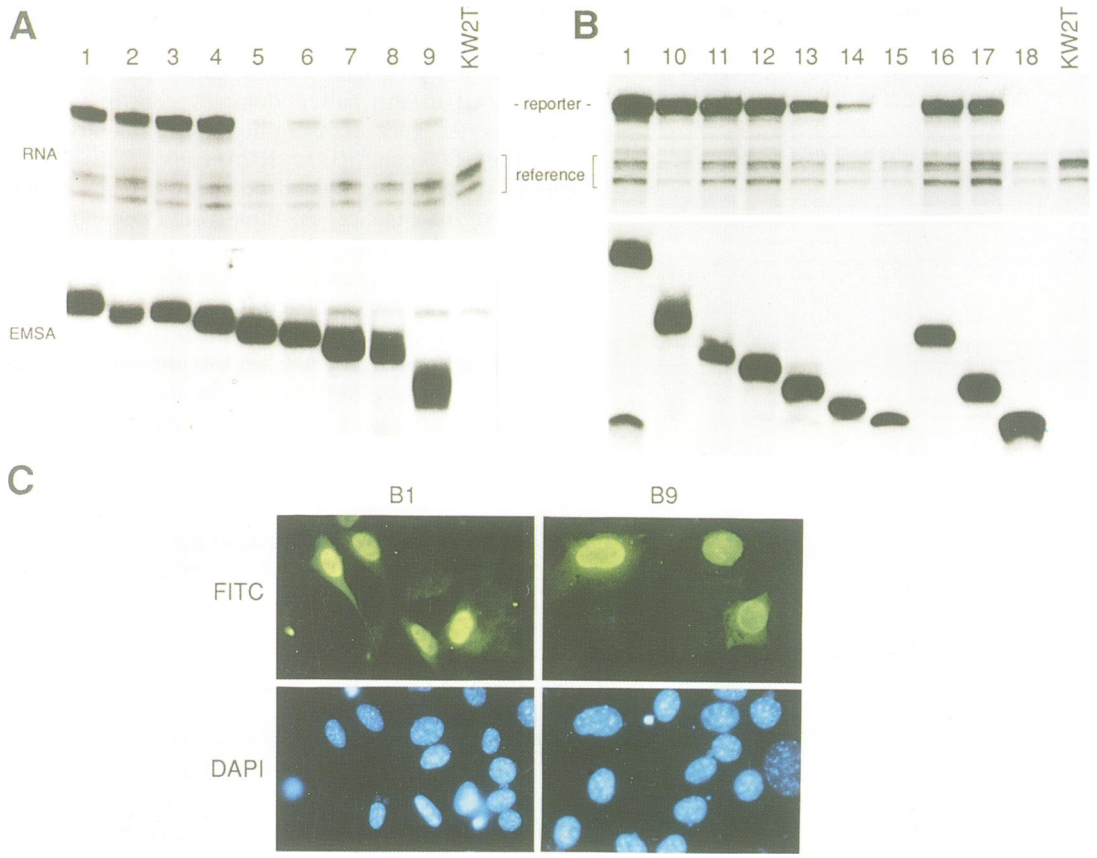

Fig. 3. Characterization of the minimal transactivation domain of BSAP by analysis of deletion mutants in SP2/0 cells. (**A**) C-terminal deletions of BSAP. Cells of the murine plasma cell line SP2/0 were transiently transfected with the glob–CD19 reporter gene (7 μg), OVEC-ref (2 μg) and expression plasmids (2 μg) directing the synthesis of full-length BSAP (B1) or the C-terminal deletion mutants B2–B9. The empty expression vector pKW2T was used as a control. Cytoplasmic RNA of transfected cells was analysed for β-globin transcripts by RNase protection assay. Protein extracts were prepared from the nuclear pellet of the same transfected cells and amounts equivalent to the mapped RNA were analysed for the presence of mutant BSAP proteins by EMSA with a labelled CD19-2(A-ins) oligonucleotide. Only the relevant part of the autoradiograph containing the protein–DNA complexes is shown. (**B**) N-terminal deletion analysis of the transactivation domain. The transcriptional activity of the internal deletion mutants B10–B17, the paired domain peptide B18 and wild-type BSAP (B1) was investigated as described in panel A. (**C**) Immunofluorescence analysis. Expression plasmids directing the synthesis of full-length BSAP (B1) and of the transcriptionally inactive mutant B9 were transiently transfected into murine COP-8 fibroblasts, and the transfected cells were stained with a polyclonal rabbit anti-BSAP paired domain antiserum (Adams *et al.*, 1992) which was detected with FITC-conjugated goat anti-rabbit antibodies. Nuclei were visualized by DAPI staining of chromosomal DNA.

transactivation domain, as determined by the use of an artificial reporter construct, is also relevant for transcription of a BSAP target gene in its natural chromatin context.

An inhibitory domain at the extreme C-terminus downregulates the transactivation function of BSAP

The transactivation function of BSAP was characterized up to this point only in transiently transfected cells of the plasma cell line SP2/0. To verify the conclusions drawn from these experiments, we measured the transactivation potential of C-terminal and internal BSAP deletion mutants in another murine plasma cell line, J558L. A representative analysis of the C-terminal deletion mutants in these cells is shown in Figure 6. To our surprise, maximal transcriptional activity of BSAP was only observed if 22 amino acids were deleted from the C-terminus in the mutant protein B3. Upon further deletion of sequences, the transactivation potential was decreased in mutant B5 to the same low level seen with the full-length protein B1. This basal level of transactivation was, however, significantly higher than the background level of β-globin transcription observed with the empty expression vector pKW2T (Figure 6) or with the paired domain peptide B18 (data not shown). Several tentative conclusions can be drawn from these results. First, the C-terminal transactivation domain of BSAP is immediately flanked by inhibitory sequences which are located at the extreme C-terminus and which appear to be inactive in SP2/0 cells. Second, the C-terminal boundary of this transactivation domain is the same in both plasma cell lines (J558L and SP2/0). Third, a weak transactivation function resides within internal BSAP sequences which are located between the C-terminal transactivation region and the N-terminal paired domain. Fourth, only the potent C-terminal transactivation sequences, but not the weakly active internal sequences, appear to be under negative control by the inhibitory domain at the extreme C-terminus.

To facilitate quantitation of the results, we repeated and extended the above transient transfection analysis with a luciferase reporter gene (luc–CD19; Figure 7E) which is under the control of the same BSAP-dependent promoter

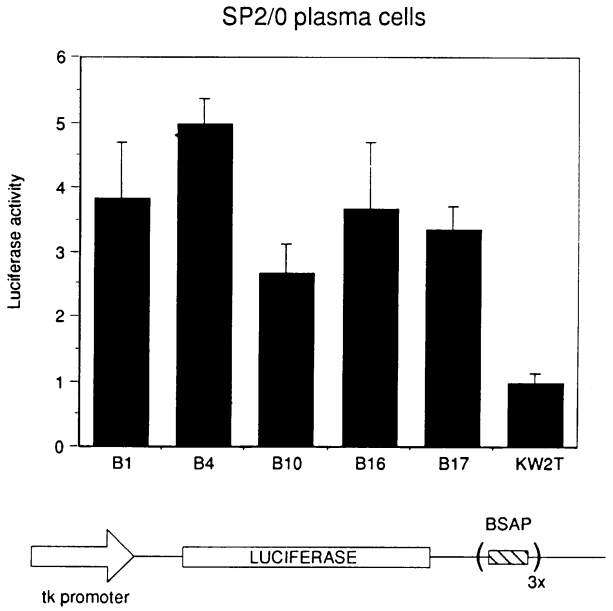

Fig. 4. The minimal transactivation domain of BSAP is sufficient to activate transcription from an enhancer position. The luciferase gene (5 μg) schematically depicted below and expression plasmids (2 μg) for wild-type BSAP (B1) or the indicated mutant proteins were transiently transfected together with a reference CMV–CAT gene (0.1 μg) into SP2/0 cells. Luciferase activities were standardized relative to CAT values to normalize for differences in transfection efficiencies. For each construct, the average luciferase value of four independent experiments is shown relative to the luciferase activity measured with the empty expression vector pKW2T. Error bars indicate standard deviations of the mean. The luciferase reporter gene used was previously described as lucTK-GAL/CD19 (Czerny and Busslinger, 1995).

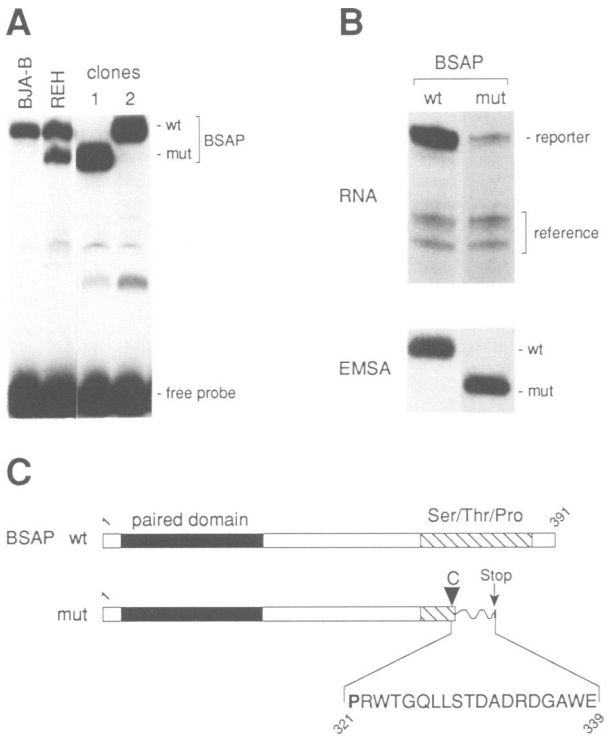

Fig. 5. REH cells express a truncated polypeptide with a frameshift mutation eliminating the transactivation function of BSAP. (**A**) EMSA analysis. Nuclear extracts of the human B-lymphoid cell lines REH and BJA-B were analysed by EMSA in parallel with whole cell extracts of COP-8 cells that were transfected with expression plasmids containing BSAP cDNA inserts isolated from REH cells by PCR amplification. A BSAP recognition sequence originating from the sea urchin H2A-2.2 gene (Barberis *et al.*, 1990) was used as binding probe. The positions of the mutant (mut) and wild-type (wt) BSAP proteins, which are expressed from the two *PAX-5* alleles represented by cDNA clones 1 and 2, are indicated. (**B**) Transactivation assay. The transcriptional activity of mutant and wild-type BSAP proteins encoded by the cDNA clones 1 and 2, respectively, was determined by transient transfection in SP2/0 cells, and the synthesis of the BSAP proteins was analysed by EMSA exactly as described in the legend of Figure 3A. (**C**) Schematic diagram of the mutated BSAP polypeptide. The insertion of a single C-residue (arrow head) generates a truncated protein due to a shift in reading frame (wavy line). This mutation affects the sequence after proline 321 of BSAP resulting in premature termination of translation after the 18 amino acid residues shown.

as the glob–CD19 gene (Figure 1). In agreement with the experiment shown in Figure 6, deletion of 33 C-terminal amino acids in mutant B4 consistently led to a 4-fold increase of the transactivation potential compared with full-length BSAP (B1). Interestingly, additional deletion of internal amino acids 170–280 in mutant B16 did not lower the maximal transcriptional activity of BSAP in J558L cells (Figure 7A). This result therefore indicates that internal sequences do not significantly contribute to the transcriptional activity of BSAP, as long as the potent C-terminal transactivation function is unmasked by deletion of sequences from the extreme C-terminus. However, deletion of the same internal region in the context of an intact C-terminus resulted in a 7-fold lower activity of the mutant protein B10 compared with full-length BSAP (B1; Figure 7A), thus corroborating that the weak transactivation function of internal sequences (amino acids 170–280) is not repressed by the C-terminal inhibitory domain in full-length BSAP (B1). In contrast, the C-terminal transactivation function is under tight negative regulation by the adjacent inhibitory domain as evidenced by a 30-fold difference in transactivation potential between mutants B10 and B16, both of which lack internal sequences (Figure 7A).

To analyse the inhibitory effect of the extreme C-terminal sequences in greater detail, we transfected increasing amounts of expression plasmids encoding the mutant proteins B10 and B16 into J558L cells and measured the transactivation potential as well as protein expression by luciferase and EMSA assays, respectively (Figure 7B). These two proteins differ from each other only by the presence (B10) or absence (B16) of the C-terminal inhibitory domain. As shown in Figure 7B, protein B10 was consistently expressed at a ~4-fold lower level than protein B16, indicating that the last 33 amino acids of BSAP may influence protein stability in J558L cells. However, comparison of the transactivation data obtained with cells expressing similar amounts of the two proteins clearly indicated that the observed discrepancies are not caused by protein instability but reflect an inherent difference in the transactivation potential of the two proteins (compare 2 μg of B16 DNA versus 8 μg of B10 DNA; Figure 7B). We conclude therefore that the potent C-terminal transactivation function of BSAP is tightly regulated in J558L cells by an adjacent inhibitory domain at the extreme C-terminus.

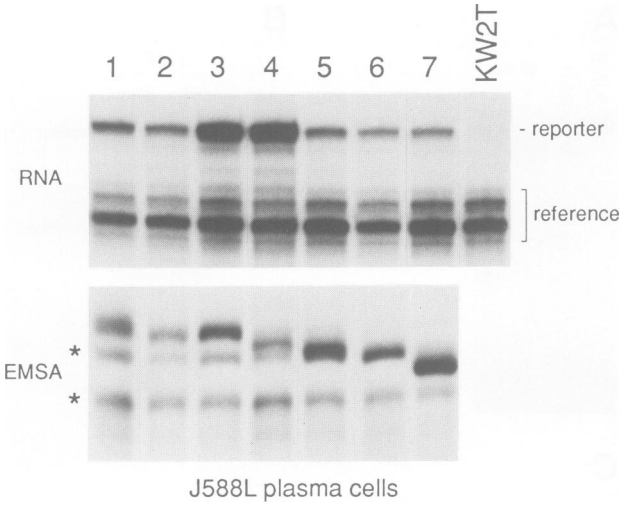

Fig. 6. Negative regulation of the C-terminal transactivation function of BSAP in J558L cells. Expression plasmids (2 μg) coding for full-length BSAP (B1) or the C-terminal deletion mutants B2–B7 were transiently transfected into cells of the murine plasma cell line J558L together with the glob–CD19 (7 μg) and OVEC–ref (2 μg) constructs. The levels of β-globin mRNA and BSAP protein were analysed by RNase protection assay and EMSA, respectively (as described in the legend to Figure 3A). Unspecific protein–DNA complexes are indicated by asterisks.

The C-terminal transactivating and inhibitory domains of BSAP function as an independent module in B cells

We next studied the function of the C-terminal region of BSAP in the more physiological background of the murine B cell line M12.4.1, which expresses endogenous BSAP. To be able to distinguish between transactivation by endogenous and exogenous BSAP, we transferred the C-terminal sequences of this transcription factor to the heterologous DNA binding domain of the yeast GAL4 protein. The GAL4–BSAP fusion proteins G1, G2 and G3 (Figure 2) were transiently expressed in M12.4.1 cells, and their transcriptional activity was determined, using a co-transfected GAL4-responsive luciferase gene as reporter system (see Figure 7E). As illustrated in Figure 7C, the chimeric protein G1 containing both the transactivating and inhibitory domains of BSAP (amino acids 281–391) was 3-fold less active than protein G2 (amino acids 281–358) lacking the inhibitory region at its C-terminus. Moreover, the transcriptional activity of mutant G3, which contained only the minimal transactivation domain (amino acids 319–358) as defined in SP2/0 cells, was almost as active as the chimeric protein G2. It is interesting to note that the transcriptional activity of protein G2 increased linearly with increasing protein amounts (Figure 7D). In contrast, the C-terminal inhibitory domain repressed the transactivation function of protein G1 to the same extent even at high protein concentrations (Figure 7D). As shown by EMSA analysis, the observed difference in transactivation potential did not result from altered expression of the mutant proteins G1 and G2 in M12.4.1 cells (Figure 7D). All these results therefore support the notion that the inhibitory domain at the extreme C-terminus of BSAP is also active in B cells and that the entire C-terminal region of BSAP functions as an independent regulatory module in the context of an unrelated DNA binding domain.

A C-terminal sequence of 55 amino acids functions as a cell type-independent transactivation domain of BSAP

During the course of these studies we realized that the plasmacytoma cell line SP2/0 was atypical in two respects. First, the inhibitory domain of BSAP appeared to be inactive under normal growth conditions in these cells. Second, the transactivation potential of BSAP was highest in SP2/0 cells compared with any other cell line tested. Both properties of the SP2/0 cells greatly facilitated the identification and characterization of the C-terminal transactivation domain of BSAP by mutational analysis. However, we also observed that the minimal transactivation domain (amino acids 319–358) as defined in SP2/0 cells did not possess maximal activity in J558L, RAC65 and M12.4.1 cells (construct G3 in Figure 7C and P.Dörfler, unpublished data). These findings suggested that amino acid residues between positions 304 and 319 of BSAP which are highly conserved among all BSAP-related Pax proteins (Figure 9) may also be required for full activity of the C-terminal transactivation domain. To test this hypothesis, we analysed the activity of chimeric GAL4–BSAP proteins containing N-terminal extensions of the BSAP transactivation domain in transiently transfected cells of the murine embryonal carcinoma cell line RAC65 (Pratt *et al.*, 1990). As shown in Figure 8, the fusion protein G4 with an N-terminal extension of the transactivation domain up to position 304 was transcriptionally as active as protein G2 (281–358). In contrast, the activity of protein G3 comprising just the minimal transactivation domain (319–358) was reproducibly lower in RAC65 cells, similar to the observations in other cell lines. Hence, only an extended sequence consisting of 55 C-terminal amino acids (304–358) functions as a potent cell type-independent transactivation domain of BSAP.

The C-terminal regulatory module consisting of transactivating and inhibitory sequences is functionally conserved in all BSAP-related Pax proteins

The mammalian BSAP (Pax-5), Pax-2 and Pax-8, the zebrafish Pax-b and the more distantly related sea urchin Pax-258 constitute a subfamily of Pax proteins with similar structural organization. All of these proteins are highly conserved in the C-terminal sequences corresponding to the transactivating and inhibitory domains of BSAP (Figure 9). The suPax-258 protein is the only notable exception as it shows little homology with the inhibitory sequences of BSAP and contains a 132 amino acid-long extension at its C-terminus which is absent in the other Pax proteins (T.Czerny and M.Busslinger, unpublished data). Given the high sequence conservation among the other Pax proteins of this subfamily, we investigated the possibility that sequence homology may also reflect functional conservation of the transactivation and inhibitory domains. We linked the GAL4 DNA binding domain either to the entire C-terminal sequence of each of the Pax proteins shown in Figure 9 (constructs '+ inhibitory domain') or only to the sequence corresponding to the transactivation domain of BSAP (constructs '– inhibitory

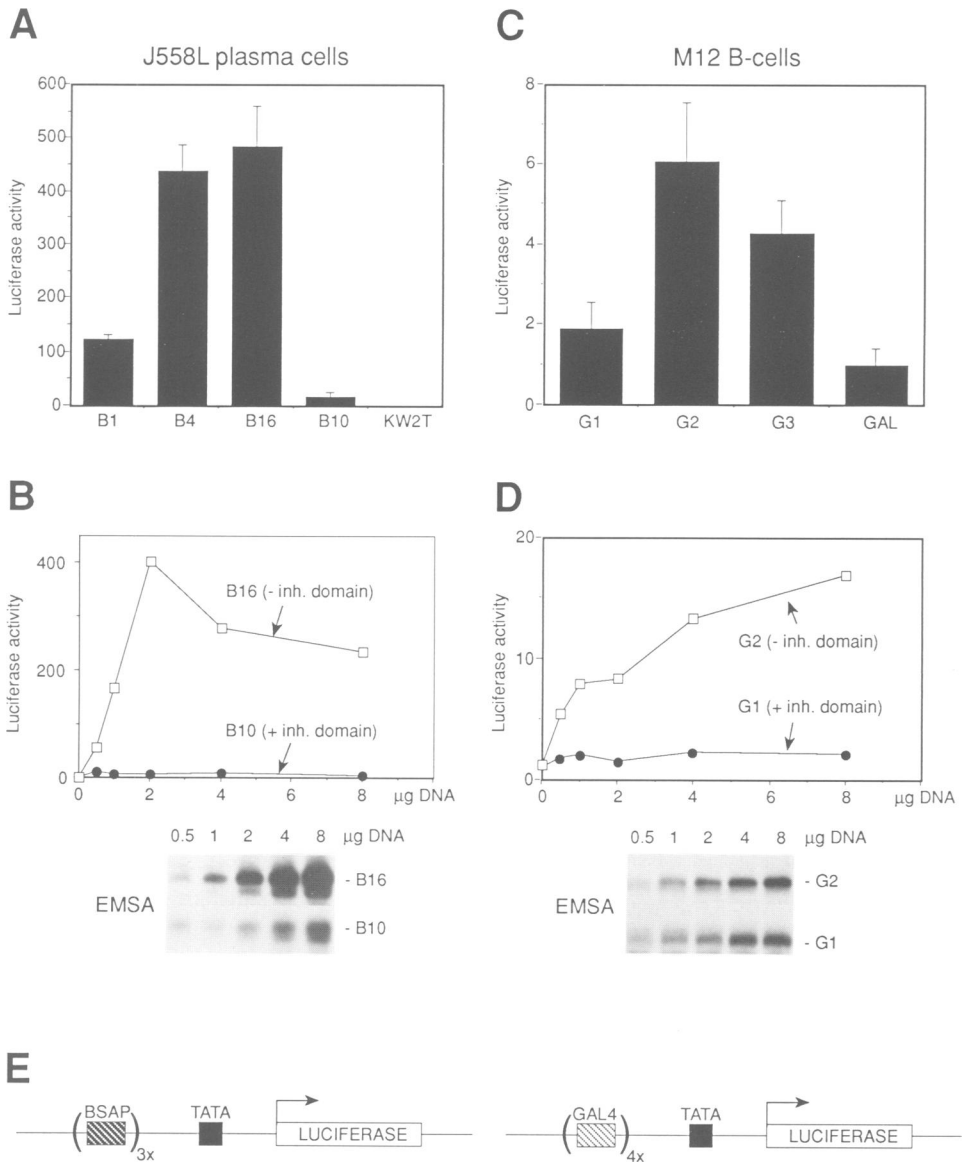

Fig. 7. An inhibitory domain at the extreme C-terminus downregulates the transactivation function of BSAP. (**A**) Transactivation potential of full-length and mutant BSAP proteins in J558L plasma cells. Expression plasmids (2 µg) directing the synthesis of wild-type (B1) or mutant (B4, B10, B16) BSAP proteins were electroporated into the plasma cell line J558L together with the luciferase gene luc–CD19 (5 µg) and the reference CMV–CAT gene (0.2 µg). Luciferase activities of three independent transfection experiments were normalized to the measured CAT activities and are shown as average values relative to the activity observed in the presence of the empty expression plasmid pKW2T. (**B**) Concentration dependence of transactivation by mutant BSAP proteins. Increasing amounts of expression vectors coding for the mutant proteins B10 and B16 were transiently transfected together with the luciferase gene luc–CD19 (5 µg) into J558L cells. The amount of expression plasmid was equalized by addition of the empty expression vector pKW2T. Luciferase activities are shown relative to the activity measured with pKW2T. The same whole cell extracts, which were used for measuring the luciferase activity, were assayed for the presence of the mutant BSAP proteins by EMSA with a labelled CD19-2(A-ins) oligonucleotide. Only the relevant part of the autoradiographs containing the protein–DNA complexes is shown. (**C**) Transactivation by chimeric GAL4–BSAP proteins in M12.4.1 B cells. Expression plasmids (1 µg) encoding the chimeric proteins G1, G2 and G3 or the DNA binding domain (GAL) of the yeast GAL4 protein were electroporated into M12.4.1 B cells together with the GAL4-responsive reporter gene plGC-luc (5 µg) and the CMV–CAT gene (0.1 µg). Normalized luciferase activities of six independent experiments are shown as average values relative to the activity measured with the GAL4 DNA binding domain alone. (**D**) Concentration-dependent transactivation by GAL4–BSAP fusion proteins. Increasing amounts of the G1 and G2 expression vectors were transfected into M12.4.1 B cells together with the reporter plGC-luc (5 µg) and reference CMV–CAT genes (0.3 µg). Luciferase activities normalized for CAT values are shown relative to the activity measured with the GAL4 DNA binding domain alone. The chimeric GAL4–BSAP proteins were detected in the whole cell extracts by EMSA with a labelled GAL4 binding site. (**E**) Schematic diagram of the luciferase reporter genes used. The BSAP-dependent reporter construct luc–CD19 (shown to the left) and the GAL4-responsive reporter gene plGC-luc (shown to the right) were previously described (Czerny and Busslinger, 1995).

domain'; Figure 10A). As shown by transient transactivation assay of these chimeric proteins in RAC65 cells (Figure 10A), all fusion proteins containing just the sequences homologous to the BSAP transactivation domain activated GAL4-dependent luciferase transcription to a much higher level than those proteins which in addition included the very C-terminal sequence related to the inhibitory domain of BSAP. Moreover, EMSA analysis with a GAL4 binding probe did not reveal significant differences in protein expression levels (Figure 10A). Taken together these data demonstrate that the regulatory module consisting of a potent transactivation region and

an adjacent inhibitory domain has been evolutionarily conserved among all members of the BSAP subfamily of Pax proteins.

The extended C-terminal sequences of the sea urchin Pax-258 protein were also capable of silencing the transactivation function, although they share little homology with the inhibitory sequences of the vertebrate Pax-2, Pax-5 and Pax-8 proteins. To localize further the inhibitory domain in the C-terminus of the suPax-258 protein, we introduced a stop codon into the GAL4–suPax-258 protein after amino acid position 427, which corresponds to the C-termini of the other Pax proteins. As shown in Figure 10B, this truncated fusion protein (338–427) was still able to fully repress the transactivation function encoded by amino acids 338 to 395 of the suPax-258 protein. Thus, the inhibitory domain of the suPax-258 protein was also mapped next to the transactivation domain in analogy to the other BSAP-related Pax proteins. As these negative regulatory sequences of suPax-258 markedly differ from the corresponding Pax-2, Pax-5 and Pax-8 sequences, it was formally possible that any arbitrary sequence at the C-terminus might serve the purpose of inhibiting the adjacent transactivation domain. To clarify this point, we deleted a single nucleotide after codon 358 in the BSAP cDNA sequence which shifted the reading frame immediately downstream of the transactivation domain, thus resulting in a totally unrelated amino acid sequence at the C-terminus of the chimeric protein G6 (Figures 2 and 9). As shown in Figure 10C, this fusion protein was transcriptionally as active in transfected RAC65 cells as protein G4 lacking the inhibitory domain. Hence, unique C-terminal sequences of the BSAP-related Pax proteins specify an inhibitory function which has been conserved together with the adjacent transactivation domain over a large evolutionary distance.

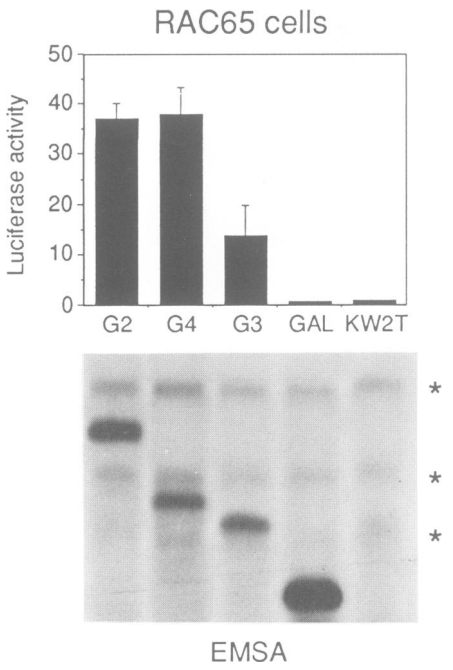

Fig. 8. Cell type-independent function of an extended BSAP transactivation domain consisting of 55 amino acids. Expression plasmids (2 μg) encoding the indicated GAL4–BSAP fusion proteins or the GAL4 DNA binding domain (GAL) were transfected into the murine embryonal carcinoma cell line RAC65 together with the reporter plGC-luc (5μg) and reference CMV–CAT (0.3 μg) genes. Normalized luciferase activities of five independent transfection experiments are shown as average values relative to the activity obtained with the empty expression vector pKW2T. The same whole cell extracts which were used for measuring luciferase activity were also assayed for the presence of the GAL4 fusion proteins by EMSA with a GAL4 binding site probe. Asterisks indicate unspecific protein–DNA complexes.

Discussion

Our analysis of the transactivation properties of BSAP (Pax-5) led to the identification of a C-terminal protein module capable of tightly controlling the activity of this transcription factor. This regulatory region consists of a potent cell type-independent transactivation domain of 55

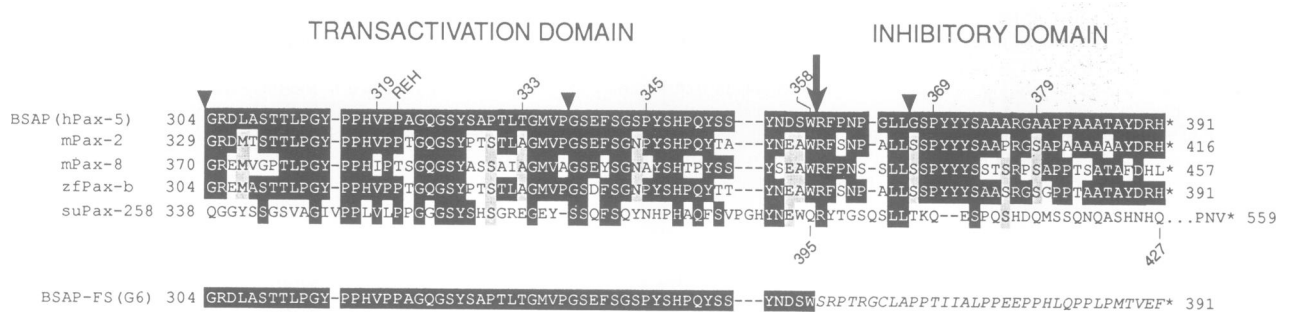

Fig. 9. Amino acid sequence comparison of the C-termini of BSAP-related Pax proteins. Amino acid sequences constituting the transactivation region and the adjacent inhibitory domain of BSAP (hPax-5) (Adams *et al.*, 1992) were aligned with the C-terminal sequences of the mouse Pax-2 (Dressler *et al.*, 1990), mouse Pax-8 (Plachov *et al.*, 1990), zebrafish Pax-b (Krauss *et al.*, 1991) and sea urchin Pax-258 proteins. The paired domain of the *Paracentrotus lividus* Pax-258 protein is highly homologous to that of the vertebrate Pax-2, Pax-5 and Pax-8 proteins, indicating that Pax-258 is a sea urchin representative of this Pax protein subfamily (T.Czerny and M.Busslinger, unpublished data). Amino acid residues that are identical to the BSAP sequence are highlighted by black overlay and homologies among at least three other members of the subfamily are indicated by grey overlay. An arrow demarcates the experimentally defined boundary separating the transactivation region and inhibitory domain. Filled arrowheads point to the position of introns in the *PAX-2*, *PAX-5* and *PAX-8* genes (Kozmik *et al.*, 1993; Sanyanusin *et al.*, 1995; Z.Kozmik and M.Busslinger, unpublished data). Numbers refer to the first and last amino acid residues of each Pax protein sequence shown. The deletion end-points of the different BSAP mutants (Figure 2) and the frameshift mutation identified in the *PAX-5* allele of REH cells (Figure 5) are shown by their respective amino acid positions above the BSAP sequence. Stop codons are denoted by asterisks and gaps introduced for optimal sequence alignment by dashed lines. The long C-terminal extension of the suPax-258 protein is symbolized by dots. The C-terminal amino acid sequence of the GAL4–BSAP protein G6 (Figure 2) which is translated in an alternate reading frame downstream of the transactivation domain due to a frameshift (FS) mutation is shown in italics.

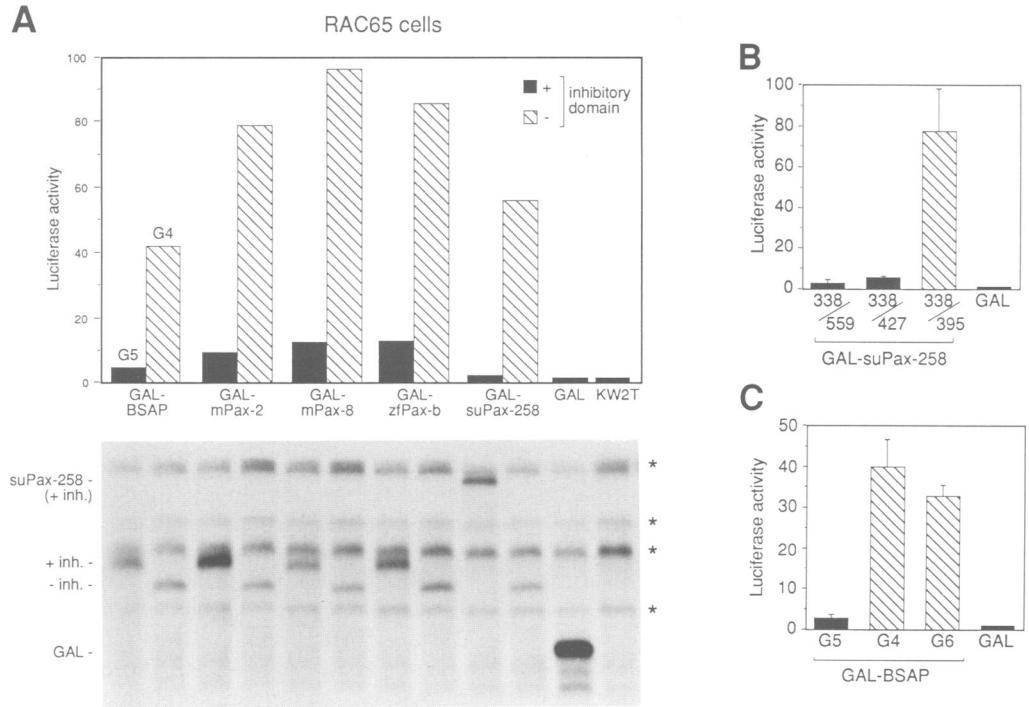

Fig. 10. The C-terminal module consisting of activating and inhibitory domains is functionally conserved in all BSAP-related Pax proteins. (**A**) Transactivation by GAL4–Pax fusion proteins. The entire C-terminal sequences of each of the Pax proteins shown in Figure 9 were fused to the DNA binding domain of GAL4 (constructs '+ inhibitory domain'). Alternatively, a stop codon was introduced into these Pax sequences at the position corresponding to the C-terminal boundary of the BSAP transactivation domain (indicated by arrow in Figure 9; constructs '– inhibitory domain'). The transactivating potential of these GAL4–Pax fusion proteins was analysed in transiently transfected RAC65 cells, and the presence of the fusion protein was detected by EMSA in whole cell extracts exactly as described in the legend of Figure 8. The positions of the different GAL4–Pax proteins are indicated to the left and unspecific protein–DNA complexes are denoted by asterisks to the right. The GAL–suPax-258 protein contained the extended C-terminal region of Pax-258 (amino acids 338–559). (**B**) Definition of the inhibitory domain of the sea urchin Pax-258 protein. GAL4 fusion proteins containing Pax-258 sequences from amino acid 338 to amino acids 395, 427 or 559, respectively, were expressed in transiently transfected RAC65, and their transactivation potential was determined as described in Figure 8. Luciferase activities of three independent experiments were normalized to the expression of the co-transfected CMV–CAT gene and are shown as average values relative to the activity measured with the GAL4 DNA binding domain (GAL) alone. (**C**) Specific sequence requirement of the inhibitory domain. The transcriptional activities of the GAL4–BSAP fusion proteins G4, G5 and G6 were measured, and the data of three independent experiments were evaluated as described in (B). The G6 protein (BSAP-FS) contained an unrelated sequence at the C-terminus (see Figure 9) instead of the inhibitory domain of BSAP.

amino acids and of adjacent inhibitory sequences at the extreme C-terminus of BSAP. The same modular arrangement of positive and negative regulatory sequences at the C-terminus was shown to be a specific feature of all members of the Pax protein subfamily consisting of the mammalian BSAP (Pax-5), Pax-2, Pax-8, zebrafish Pax-b and sea urchin Pax-258 proteins. The C-terminal regions of most Pax proteins including the BSAP-related transcription factors are characterized by a preponderance of serine, threonine and proline residues. However, these C-terminal regions vastly differ in primary sequence between members of individual Pax protein subfamilies (Walther *et al.*, 1991). To date, potent transactivation functions have been assigned to the C-terminal sequences of Pax-3 (Chalepakis *et al.*, 1994) and Pax-6 (Glaser *et al.*, 1994; Czerny and Busslinger, 1995). It appears unlikely that the transactivation domains of these two Pax proteins are also flanked by negative regulatory sequences analogous to those of the BSAP-related Pax transcription factors. First, the C-terminal 143 amino acids of the mouse Pax-6 protein endow a chimeric GAL4–mPax-6 fusion protein with high transcriptional activity in J558L cells (Czerny and Busslinger, 1995) despite the fact that these cells show maximal silencing of the transactivation function of BSAP. Second, functional analysis has recently localized a potent transactivation domain within the last 78 C-terminal amino acids of the Pax-3 protein, while an inhibitory activity was mapped to the first 90 N-terminal residues of this transcription factor (Chalepakis *et al.*, 1994).

The C-terminal regulatory module of BSAP-related Pax proteins is highly biased in its amino acid composition. Its 88 amino acid-long sequence contains 11 or more proline residues and >25 hydroxylated amino acids (serine, threonine and tyrosine). As a consequence, the entire domain is predicted to be exposed as an unstructured and hydrophilic tail on the surface of these Pax proteins. It is interesting to note that the entire C-terminal module is encoded by three exons with conserved intron positions in the *PAX-2* (Sanyanusin *et al.*, 1995), *PAX-8* (Kozmik *et al.*, 1993) and *PAX-5* genes (Busslinger *et al.*, 1996; see Figure 9). The first and last of these exons code exclusively for transactivating or inhibitory amino acid residues, respectively, while the second exon contributes to both protein domains. Amino acid 358 of BSAP was experimentally determined to be at the boundary between the transactivating and inhibitory sequences (see Figure 9). It is, however, conceivable that the negative regulatory

domain extends N-terminally up to the next exon–intron boundary (amino acid 338) and may thus partially overlap with the transactivation region (Figure 9). A possible N-terminal extension would also increase the homology of the inhibitory sequences of the sea urchin Pax-258 protein with those of the other BSAP-related Pax proteins. In support of such an idea, the sequences from amino acids 359 to 391 of BSAP failed to function as an autonomous inhibitory domain upon transplantation to a heterologous protein, as they were unable to silence the acidic and serine/threonine-rich transactivation domains of GAL4–VP16 and GAL4–TFE3 (Seipel *et al.*, 1992), respectively (P.Dörfler, unpublished data). Alternatively, the negative regulatory domain could act in a sequence-specific manner to suppress only the transactivation function of BSAP-related Pax proteins in analogy to a recently described inhibitory domain of c-Fos which is able to inactivate just one (HOB1) of several transactivation modules (Brown *et al.*, 1995).

The few target genes which are to date known to be regulated by BSAP-related Pax proteins indicate that critical recognition sequences for these transcription factors are found in various positions relative to the site of transcription initiation. Functional binding sites for Pax-5 (BSAP) have been identified in a midbrain-specific enhancer of the *Engrailed-2* gene (Song *et al.*, 1996) and in the downstream 3′α enhancer of the immunoglobulin heavy chain gene locus (Singh and Birshtein, 1993; Neurath *et al.*, 1994). In addition, recognition sequences for the sea urchin Pax-258 and mammalian Pax-8 proteins were found upstream of the TATA box in the proximal promoters of a subset of sea urchin histone genes (Barberis *et al.*, 1989) and the mammalian thyroperoxidase gene (Zannini *et al.*, 1992), respectively. In our transactivation study we have used artificial reporter genes containing Pax binding sites in proximal promoter and distal enhancer positions, thus approximating the situation described for the above target genes. It is therefore reasonable to assume that transcription of these natural target genes is dependent on the C-terminal regulatory module of the respective Pax protein similar to the expression of the reporter genes analysed. The situation is, however, different for the BSAP target genes *CD19* and *Iε* coding for a B cell surface protein and the germline immunoglobulin Cε transcript, respectively. Both of these genes contain a high-affinity BSAP binding site instead of a TATA box in the –30 region upstream of a cluster of heterogeneous transcription start sites (Rothman *et al.*, 1991; Kozmik *et al.*, 1992). Hence it is conceivable that BSAP regulates transcription initiation of these genes through direct contact with the basal transcription machinery which may involve BSAP domains other than the C-terminal regulatory module. In this context it is interesting to note that we have identified a frameshift mutation in one *PAX-5* allele of the human ALL cell line REH which shifts the reading frame immediately upstream of the minimal transactivation domain of BSAP. This cell line therefore expresses a putative transdominant protein from the mutated allele and only half of the normal amount of transcriptionally competent BSAP from the remaining wild-type allele. REH cells express considerably lower *CD19* mRNA levels than other established B-lymphoid cell lines (Kozmik *et al.*, 1992). Moreover, we have recently discovered that *CD19* expression is exquisitely sensitive to the dosage of BSAP protein in cultured murine pro-B cells (S.Nutt and M.Busslinger, unpublished data). The lower expression level of the endogenous *CD19* gene in REH cells could therefore be a direct consequence of truncating the BSAP protein by frameshift mutation, suggesting that intact C-terminal sequences of BSAP are also required for efficient transcription of the TATA-less *CD19* gene.

Silencing of a transactivation domain by inhibitory sequences has been described for several other transcription factors, and different mechanisms for this negative regulation have been proposed. In the case of the c-Jun protein, a cell type-specific inhibitory protein was shown to interact with negative regulatory sequences (δ and ε domains) which flank and thus mask the N-terminal transactivation domain a1 (Baichwal *et al.*, 1992). Likewise, a cellular protein has been implicated in silencing the transactivation function of c-Fos (Brown *et al.*, 1995) and Sp1 (Murata *et al.*, 1994) by binding to inhibitory sequences located adjacent to a transactivation domain. The yeast bHLH protein PHO4 represents the best studied case of a transcription factor whose transactivation domain is masked by protein–protein interaction. Inhibitory sequences present on either side of the transactivation domain of PHO4 are responsible for binding of the cyclin-like repressor PHO80 (Jayaraman *et al.*, 1994). A different, promoter-specific repression mechanism has been postulated for the transcription factor C/EBPα which is involved in the regulation of the albumin gene. The function of an inhibitory domain which overlaps with one of three transactivation region of C/EBPα was alleviated only in the cooperative environment of the albumin promoter (Nerlov and Ziff, 1994). Finally, the transactivation domains of C/EBPβ (NF-M) (Kowenz-Leutz *et al.*, 1994), c-Myb (Dubendorff *et al.*, 1992) and the mammalian heat shock transcription factor 1 (Shi *et al.*, 1995; Zuo *et al.*, 1995) appear to be masked from contacting the transcription machinery by intramolecular interaction of inhibitory and transactivating sequences. Whether any of these negative regulatory mechanisms also applies to the BSAP-related Pax protein, remains to be seen. It is, however, worth mentioning that BSAP has been implicated in the repression of the 3′α enhancer of the immunoglobulin heavy chain gene locus (Singh and Birshtein, 1993; Neurath *et al.*, 1994) which could be mediated by the C-terminal inhibitory domain of BSAP in the regulatory context of this enhancer.

Negatively regulated transactivation domains appear to be a characteristic feature of inducible transcription factors which are activated in response to intracellular signalling. This is true for almost all of the transcriptional regulators discussed above. One of the best studied examples is the transcription factor C/EBPβ (NF-M) which has been implicated as a nuclear target for cytokine-induced gene expression in different cell types (Kowenz-Leutz *et al.*, 1994). C/EBPβ (NF-M) contains two inhibitory domains (CR5 and CR7) which interact *in cis* with the transactivation domain and thus silence its function. A highly conserved MAP kinase site in one (CR7) of these inhibitory regions is directly involved in the negative interaction which masks the transactivation domain. Phosphorylation of this site in response to activation of MAP kinase

pathways disrupts this intramolecular interaction and thus results in derepression of the constitutive transactivation domain of C/EBPβ (NF-M) (Kowenz-Leutz *et al.*, 1994). The presence of a negatively regulated transactivation domain in Pax-2, Pax-5 and Pax-8 also strongly suggests that these transcription factors are at the receiving end of signal transduction in the nucleus. However, all our efforts to demonstrate a link between transcriptional activation of BSAP and signal transduction have so far failed. Stimulation of B cell lines with various interleukins, serum growth factors or phorbol esters did not inactivate the silencing function of the C-terminal inhibitory domain (P.Dörfler, unpublished data). Thus it will be a challenge for the future to identify the appropriate signals and the molecular mechanisms through which BSAP and its related transcription factors are relieved from negative regulation by their C-terminal inhibitory domain.

Materials and methods

Cell lines

The murine B cells line M12.4.1 (Glimcher *et al.*, 1982) and the human B-lymphoid cell lines REH (Rosenfeld *et al.*, 1977) and BJA-B (Barberis *et al.*, 1990) all express endogenous BSAP. In contrast, the murine plasmacytoma cell lines SP2/0 and J558L (Barberis *et al.*, 1990), the embryonal carcinoma cell line RAC65 (Pratt *et al.*, 1990) and the murine fibroblast cell line COP-8 (Tyndall *et al.*, 1981) do not express any endogenous BSAP. The cell lines SP2/0, J558L, RAC65 and COP-8 were grown in high glucose DMEM medium supplemented with 10% fetal calf serum, while M12.4.1, BJA-B and REH cells were cultured in RPMI 1640 medium containing 10% fetal calf serum.

Plasmid constructions

The reporter constructs glob–CD19 and glob–CD19mut were created by cloning three copies of either the high-affinity BSAP binding site CD19-2(A-ins) (Czerny *et al.*, 1993) or a mutated site (oligonucleotides 1/2; lacking the A-insertion and containing the G to T mutation underlined) into the *Sal*I site of OVEC-S (Westin *et al.*, 1987). The SV40 enhancer was cloned into the *Eco*RI site downstream of the β-globin gene.

The eukaryotic expression vector pKW2T is derived from pRK7 (R.Klein and D.Goeddel, unpublished) by the same modifications described for the construction of pKW10 (Adams *et al.*, 1992). In addition, the double-stranded oligonucleotide 3/4 containing stop codons in all reading frames was inserted downstream of the polylinker into the *Eco*RI and *Cla*I sites.

The BSAP expression plasmids B1, B2, B4, B6, B8, B9 and B18 were generated by cloning C-terminally resected BSAP cDNAs (by exonuclease III digestion; Adams *et al.*, 1992) into the *Hin*dIII and *Eco*RI sites of pKW2T. Additional C-terminal truncations of BSAP cDNA were created by PCR using oligonucleotide 5 as common 5′ primer and oligonucleotides 6 (B3), 7 (B5) or 8 (B7) as 3′ primers, respectively, followed by cloning of the *Bgl*II–*Hin*dIII cDNA fragments into the *Bam*HI and *Hin*dIII sites of pKW10 (Adams *et al.*, 1992). For plasmids B10–B17, PCR fragments with *Eco*RI ends were generated and inserted into the *Eco*RI site downstream of the paired domain of clone B18 (except for clone B13, where the cDNA fragment was obtained by N-terminal exonuclease III digestion). The following primer pairs were used for generating these BSAP cDNA inserts: 9/10 (B10); 11/15 (B11); 12/15 (B12); 13/15 (B14); 14/15 (B15) using full-length BSAP cDNA as template and 9/16 (B16) and 12/16 (B17) using B4 cDNA as template.

The expression plasmid pKW-GAL-147 (referred to as GAL in Figure 2) was constructed by cloning the DNA binding domain (amino acids 1–147) of GAL4 as a 460 bp *Hin*dIII–*Cla*I fragment from pGal4-147 (Bergers *et al.*, 1995) into the *Hin*dIII and *Eco*RI sites of pKW2T. Ligation of the blunt-ended *Cla*I and *Eco*RI sites restored the downstream *Eco*RI site used for insertion of Pax cDNA fragments. The GAL4–BSAP fusions G1, G2, G3, G5 and G4 were generated by inserting the same PCR fragments used for constructing B10, B16, B17, B11 and the additional *Eco*RI PCR fragment generated with primers 11/16 on template B4 into the *Eco*RI site of pKW-GAL-147. Analogous GAL4 fusion constructs were generated with PCR fragments obtained from mPax-2a cDNA (Kozmik *et al.*, 1993) with primers 17/18 (long insert) and 17/19 (short insert), from mPax-8a cDNA (Kozmik *et al.*, 1993) with primers 20/21 (long) and 20/22 (short), from zfPax-b cDNA (Krauss *et al.*, 1991) with primers 23/24 (long) and 23/25 (short), and from suPax-258 cDNA with primers 26/27 (long), 26/28 (short) and 29/30 (middle; cloned as *Cla*I–*Bam*HI PCR fragment into the *Cla*I and *Bam*HI sites of pGal4-147). The BSAP cDNA part of the GAL4–BSAP frameshift mutant G6 was assembled by ligating two *Eco*RI–*Xba*I fragments, obtained by PCR amplification with primers 11/31 and 32/33, into the *Eco*RI site of pKW-GAL-147.

PCR cloning of mutated BSAP from REH cDNA

Poly(A)$^+$ RNA (4 μg) from the cell line REH (Rosenfeld *et al.*, 1977) was reverse-transcribed into cDNA by priming with p(dT)$_{15}$, and BSAP encoding sequences were isolated by PCR amplification with primers 5/34 as described (Adams *et al.*, 1992). The amplified *Hin*dIII–*Bgl*II fragments were cloned into the *Hin*dIII and *Bam*HI sites of the pKW10 expression vector (Adams *et al.*, 1992). The cDNA inserts of selected clones coding for truncated or full-length BSAP were analysed by DNA sequencing.

Oligonucleotides

The following oligonucleotides were used in this study: 1, 5′-TCGACCTGGAGAATGGGGCCTGAGGCGTTACCACCGCCTTCCTCTCTGG-3′; 2, 5′-TCGACCAGAGAGGAAGGCGGTGGTAACGCCTCAGGCCCCATTCTCCAGG-3′; 3, 5′-AATTCTAATTAACTAGAT-3′; 4, 5′-CGATCTAGTTAATTAG-3′; 5, 5′-GCGAGATCTCCATTTCATCAAGTCCTGAAA-3′; 6, 5′-GAGAGAAGCTTGCGGCCGCTAGGGGGAGCCAAGCAGCCCCG-3′; 7, 5′-GAGAGAAGCTTGCGGCCGCTAGGGACTCCCGGAAAACTCA-3′; 8, 5′-GAGAGAAGCTTGCGGCCGCTAGGGGACGTGTGGAGGGTACCCG-3′; 9, 5′-GGCGAATTCCCCCACCCCTGCTGACATCGGG-3′; 10, 5′-GCGGAATTCGGTGCCATCAGTGTTTGGTGCC-3′; 11, 5′-GGCGAATTCCGGCCGTGACTTGGCGAGCACG-3′; 12, 5′-GGCGAATTCCGTCCCCCCCCGCTGGACAGGG-3′; 13, 5′-GGCGAATTCCCCCTACAGCCACCCTCAGTAT-3′; 14, 5′-GGCGAATTCCTGGAGGTTCCCCAACCCG-3′; 15, 5′-GCGGAATTCGGTGCCATCAGTGTTTGGTGCC-3′; 16, 5′-GTAACCATTATAAGCTGC-3′; 17, 5′-GGCGAATTCCGGTCGTGATATGACGAGCA-3′; 18, 5′-GGCGAATTCGCCTGAAGCTTGATGTGGTCC-3′; 19, 5′-GGCGAATTCGCTCCAAGCCTCATTGTAGGCG-3′; 20, 5′-GGCGAATTCAGGGCGAGAGATGGTGGGG-3′; 21, 5′-GGCGAATTCACTACAGATGGTCAAAGGCTG-3′; 22, 5′-GGCGAATTCGCTCCAGGCCTCACTGTAGGAGGA-3′; 23, 5′-GGCGAATTCAGGTCGAGAGATGGCGAGCA-3′; 24, 5′-GGCGAATTCGCTGTGGTCCTGCAGTTTGAC-3′; 25, 5′-GGCGAATTCGCTCCAAGCTTCATTGTAAGTTGT-3′; 26, 5′-GGCGAATTCTCAAGGCGGTTACTCCAGCG-3′; 27, 5′-GGCGAATTCAGGTTCATGATGTTGGTGGTA-3′; 28, 5′-GGCGAATTCGCTTTGCCATTCGTTATAGTGCCC-3′; 29, 5′-GAGATCGATTGGCGGTTACTCCAGCGGTTCGG-3′; 30, 5′-GAGGGATCCTAATGGTTGTGACTGGCCTGATTT-3′; 31, 5′-GGCTCTAGACCAGGAGTCGTTGTACGAGGA-3′; 32, 5′-GGCTCTAGACCAACCCGGGGCTGCTTGGCT-3′; 33, 5′-GGCGAATTCGGTCATAGGCAGTGGCGGCT-3′; 34, 5′-GCGAAGCTTGGTGCCATCAGTGTTTGGTGCC-3′; 35, 5′-GATCTCCGGAGGACAGTCCTCCGATGTCCG-3′; 36, 5′-GATCCGGACATCGGAGGACTGTCCTCCGGA-3′.

Transient transfection, luciferase and CAT assays

The B cell line M12.4.1 and the plasma cells SP2/0 and J558L were transiently transfected by electroporation. Typically 2×10^7 cells were incubated with DNA (amounts indicated in figure legends) in 400 μl PBS for 5 min at room temperature in a cuvette with an interelectrode distance of 0.4 cm and then electroporated with a Bio-Rad gene pulser set at 350 V and 500 μF. After incubation for 15–20 min at 37°C, cells were transferred to 15 ml of culture medium and grown for 40 h at 37°C. RAC65 cells were transfected by calcium phosphate co-precipitation.

For luciferase and CAT assays, cells were washed twice with PBS, resuspended in 100 μl of 0.25 M Tris buffer (pH 7.5) and lysed by three freeze–thaw cycles. Thirty microliters of each supernatant were used for measuring the luciferase activity in a CliniLumat bioluminescence counter (Berthold, Wildbach, Germany) and for determining the CAT protein concentration by ELISA (CAT-ELISA kit, Boehringer Mannheim), respectively. Where indicated, these extracts were also used for EMSA analysis (4 μl per reaction).

RNA preparation and RNase protection assay

Cytoplasmic RNA was prepared according to Favaloro *et al.* (1980). The nuclear pellets which were obtained as a side product of this method

were used for preparation of nuclear extracts (see below). RNase protection assays were performed according to Vitelli *et al.* (1988) except that a hybridization temperature of 60°C (instead of 65°C) and a rabbit β-globin riboprobe (Barberis *et al.*, 1990) were used.

Nuclear extract preparation and electrophoretic mobility shift assay

The nuclei obtained as pellet during the preparation of cytoplasmic RNA from transfected cells (see above) were extracted for 15 min at 4°C with 50 µl of a buffer containing 420 mM NaCl, 20 mM HEPES (pH 7.9), 25% glycerol, 1 mM EDTA, 1 mM DTT and a protease inhibitor cocktail previously described (Adams *et al.*, 1992). Equal protein amounts of these nuclear extracts or, where indicated, of whole cell extracts obtained by freeze–thaw lysis (see above) were analysed by EMSAs according to Barberis *et al.* (1990) except for the further addition of 20 µg BSA to the binding reaction. EMSAs with GAL4–Pax fusion proteins were carried out by including 6 mM $MgCl_2$ in the binding reaction and by prolonging the incubation time to 20 min. End-labelled oligonucleotides containing the Pax binding sites of the H2A-2.2 and CD19 genes (Czerny *et al.*, 1993) were used as DNA probes to detect BSAP proteins and the double-stranded oligonucleotide 35/36 to detect GAL4 fusion proteins.

Indirect immunofluorescence assay

COP-8 cells were transiently transfected with 4 µg of BSAP expression plasmids as previously described (Adams *et al.*, 1992). Two days after transfection, cells were directly fixed on the plate by incubation with 3.7% formaldehyde in PBS for 15 min at room temperature (RT) and permeabilized with 0.5% NP-40, 0.5% BSA in PBS for 15 min at RT. Circular areas of ~1.5 cm diameter were marked on the plate with a hydrophobic pen, and cells within this area were incubated with a rabbit polyclonal serum directed against the BSAP paired domain (Adams *et al.*, 1992) followed by detection with FITC-conjugated goat anti-rabbit antibodies (each incubations for 30 min at 37°C in a moist environment). Antibody dilutions and washing steps were performed with 0.5% BSA in PBS. The stained cells were embedded in Elvanol containing 0.5 µg/ml of DAPI and photographed in a fluorescence microscope.

Acknowledgements

We are grateful to Thomas Czerny for providing the sea urchin Pax-258 cDNA, to Anders Fjose for the zebrafish Pax-b cDNA, to G.Schaffner for oligonucleotide synthesis, to I.Botto and R.Kurzbauer for DNA sequencing, to H.Tkadletz for graphical work and to M.Cotten and T.Jenuwein for critical reading of the manuscript. This work was, in part, supported by a grant from the Austrian Industrial Research Promotion Fund.

References

Adams,B., Dörfler,P., Aguzzi,A., Kozmik,Z., Urbánek,P., Maurer-Fogy,I. and Busslinger,M. (1992) *Genes Dev.*, **6**, 1589–1607.

Baichwal,V.R., Park,A. and Tjian,R. (1992) *Genes Dev.*, **6**, 1493–1502.

Barberis,A., Superti-Furga,G., Vitelli,L., Kemler,I. and Busslinger,M. (1989) *Genes Dev.*, **3**, 663–675.

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

Bergers,G., Graninger,P., Braselmann,S., Wrighton,C. and Busslinger,M. (1995) *Mol. Cell. Biol.*, **15**, 3748–3758.

Bopp,D., Jamet,E., Baumgartner,S., Burri,M. and Noll,M. (1989) *EMBO J.*, **8**, 3447–3457.

Brown,H.J., Sutherland,J.A., Cook,A., Bannister,A.J. and Kouzarides,T. (1995) *EMBO J.*, **14**, 124–131.

Busslinger,M. and Urbánek,P. (1995) *Curr. Opin. Genet. Dev.*, **5**, 595–601.

Busslinger,M., Klix,N., Pfeffer,P., Graninger,P.G. and Kozmik,Z. (1996) *Proc. Natl Acad. Sci. USA*, in press.

Chalepakis,G., Fritsch,R., Fickenscher,H., Deutsch,U., Goulding,M. and Gruss,P. (1991) *Cell*, **66**, 873–884.

Chalepakis,G., Jones,F.S., Edelman,G.M. and Gruss,P. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 12745–12749.

Czerny,T. and Busslinger,M. (1995) *Mol. Cell. Biol.*, **15**, 2858–2871.

Czerny,T., Schaffner,G. and Busslinger,M. (1993) *Genes Dev.*, **7**, 2048–2061.

Dressler,G.R., Deutsch,U., Chowdhury,K., Nornes,H.O. and Gruss,P. (1990) *Development*, **109**, 787–795.

Dubendorff,J.W., Whittaker,L.J., Eltman,J.T. and Lipsick,J.S. (1992) *Genes Dev.*, **6**, 2524–2535.

Favaloro,J., Treisman,R. and Kamen,R. (1980) *Methods Enzymol.*, **65**, 718–749.

Fickenscher,H.R., Chalepakis,G. and Gruss,P. (1993) *DNA Cell Biol.*, **12**, 381–391.

Glaser,T., Jepeal,L., Edwards,J.G., Young,S.R., Favour,J. and Maas,R.L. (1994) *Nature Genet.*, **7**, 463–471.

Glimcher,L.H., Hamano,T., Asofsky,R., Heber-Katz,E., Schwartz,R.H. and Paul,W.E. (1982) *Nature*, **298**, 283–286.

Jayaraman,P.-S., Hirst,K. and Goding,C.R. (1994) *EMBO J.*, **13**, 2192–2199.

Kowenz-Leutz,E., Twamley,G., Ansieau,S. and Leutz,A. (1994) *Genes Dev.*, **8**, 2781–2791.

Kozmik,Z., Wang,S., Dörfler,P., Adams,B. and Busslinger,M. (1992) *Mol. Cell. Biol.*, **12**, 2662–2672.

Kozmik,Z., Kurzbauer,R., Dörfler,P. and Busslinger,M. (1993) *Mol. Cell. Biol.*, **13**, 6024–6035.

Krauss,S., Johansen,T., Korzh,V. and Fjose,A. (1991) *Development*, **113**, 1193–1206.

Murata,Y., Kim,H.G., Rogers,K.T., Udvadia,A.J. and Horowitz,J.M. (1994) *J. Biol. Chem.*, **269**, 20674–20681.

Nerlov,C. and Ziff,E.B. (1994) *Genes Dev.*, **8**, 350–362.

Neurath,M.F., Strober,W. and Wakatsuki,Y. (1994) *J. Immunol.*, **153**, 730–742.

Noll,M. (1993) *Curr. Opin. Genet. Dev.*, **3**, 595–605.

Plachov,D., Chowdhury,K., Walther,C., Simon,D., Guenet,J.L. and Gruss,P. (1990) *Development*, **110**, 643–651.

Pratt,M.A.C., Kralova,J. and McBurney,M.W. (1990) *Mol. Cell. Biol.*, **10**, 6445–6453.

Rosenfeld,C., Goutner,A., Choquet,C., Venuat,A.M., Kayibanda,B., Pico,J.L. and Greaves,M.F. (1977) *Nature*, **267**, 841–843.

Rothman,P., Li,S.C., Gorham,B., Glimcher,L., Alt,F. and Boothby,M. (1991) *Mol. Cell. Biol.*, **11**, 5551–5561.

Sanyanusin,P., Schimmenti,L.A., McNoe,L.A., Ward,T.A., Pierpont, M.E.M., Sullivan,M.J., Dobyns,W.B. and Eccles,M.R. (1995) *Nature Genet.*, **9**, 358–364.

Schäfer,B.W., Czerny,T., Bernasconi,M., Genini,M. and Busslinger,M. (1994) *Nucleic Acids Res.*, **22**, 4574–4582.

Seipel,K., Georgiev,O. and Schaffner,W. (1992) *EMBO J.*, **11**, 4961–4968.

Shi,Y., Kroeger,P.E. and Morimoto,R.I. (1995) *Mol. Cell. Biol.*, **15**, 4309–4318.

Singh,M. and Birshtein,B.K. (1993) *Mol. Cell. Biol.*, **13**, 3611–3622.

Song,D.-L., Chalepakis,G., Gruss,P. and Joyner,A.L. (1996) *Development*, **122**, 627–635.

Stuart,E.T., Kioussi,C., and Gruss,P. (1994) *Annu. Rev. Genet.*, **28**, 219–236.

Treisman,J., Harris,E. and Desplan,C. (1991) *Genes Dev.*, **5**, 594–604.

Tyndall,C., La,M.G., Thacker,C.M., Favaloro,J. and Kamen,R. (1981) *Nucleic Acids Res.*, **9**, 6231–6250.

Urbánek,P., Wang,Z.-Q., Fetka,I., Wagner,E.F. and Busslinger,M. (1994) *Cell*, **79**, 901–912.

Vitelli,L., Kemler,I., Lauber,B., Birnstiel,M.L. and Busslinger,M. (1988) *Dev. Biol.*, **127**, 54–63.

Walther,C., Guenet,J.L., Simon,D., Deutsch,U., Jostes,B., Goulding, M.D., Plachov,D., Balling,R. and Gruss,P. (1991) *Genomics*, **11**, 424–434.

Westin,G., Gerster,T., Müller,M.M., Schaffner,G. and Schaffner,W. (1987) *Nucleic Acids Res.*, **15**, 6787–6798.

Zannini,M., Francis-Lang,H., Plachov,D. and Di Lauro,R. (1992) *Mol. Cell. Biol.*, **12**, 4230–4241.

Zuo,J., Rungger,D. and Voellmy,R. (1995) *Mol. Cell. Biol.*, **15**, 4319–4330.

Received on October 25, 1995; revised on December 1, 1995