The EMBO Journal vol.8 no.9 pp.2543 - 2550, 1989

A family of octamer-specific proteins present during mouse embryogenesis: evidence for germline-specific expression of an Oct factor

Hans R.Schöler, Antonis K.Hatzopoulos, Rudi Balling, Noriaki Suzuki and Peter Gruss

Max-Planck Institute for Biophysical Chemistry, Department of Molecular Cell Biology, 3400 Göttingen, FRG

Communicated by P.Gruss

We have analysed various adult organs and different developmental stages of mouse embryos for the presence of octamer-binding proteins. A variety of new octamer-binding proteins were identified in addition to the previously described Oct1 and Oct2. Oct1 is ubiquitously present in murine tissues, in agreement with cell culture data. Although Oct2 has been described as a B-cell-specific protein, similar complexes were also found with extracts from brain, kidney, embryo and sperm. In embryo and brain at least two other proteins, Oct3 and Oct7, are present. A new microextraction procedure allowed the detection of two maternally expressed octamer-binding proteins, Oct4 and Oct5. Both proteins are present in unfertilized oocytes and embryonic stem cells, the latter containing an additional protein, Oct6. Whereas Oct4 was not found in sperm or testis, it is expressed in male and female primordial germ cells. Therefore Oct4 expression is specific for the female germline at later stages of germ cell development. Our results indicate that a family of octamer-binding proteins is present during mouse development and is differentially expressed during early embryogenesis. Protease clipping experiments of Oct4 and Oct1 suggest that both proteins contain similar DNA-binding domains.

Key words: germline/embryogenesis/enhancer/maternal/octamer

Introduction

Transcriptional regulation depends on the sequence-specific interaction of *trans*-acting proteins with *cis*-acting DNA elements (Hatzopoulos *et al.*, 1988). The octamer motif is a well-characterized *cis*-acting regulatory sequence found in many promoters and enhancers (for refs see Landolfi *et al.*, 1986; Hatzopoulos *et al.*, 1988; Schöler *et al.*, 1988). This transcriptional element is necessary and sufficient for lymphoid-specific expression of the immunoglobulin genes (Dreyfus *et al.*, 1987; Wirth *et al.*, 1987). The same motif is required for ubiquitous expression of snRNA and histone genes (Tanaka *et al.*, 1988 and refs therein). Insight into this apparent paradox was provided when two proteins interacting with this sequence, referred to here as Oct1 and Oct2, were characterized (Singh *et al.*, 1985; Staudt *et al.*, 1986; Rosales *et al.*, 1987). In all cell types tested, Oct1 was present, while Oct2 was detected only in B-lymphocytes. Apparently Oct1 regulates the ubiquitously expressed genes, and Oct2 the lymphoid-specific genes (Fletcher *et al.*, 1987; Scheidereit *et al.*, 1987). However, the fact that oligomers of the octamer motif were active only in B-cells (Gerster *et al.*, 1987), indicates that additional *cis*-acting elements and/or *trans*-acting proteins are required for the function of Oct1 (Tanaka *et al.*, 1988).

Sequence analysis of Oct1 and Oct2 cDNAs revealed the presence of a homeobox (Clerc *et al.*, 1988; Ko *et al.*, 1988; Müller *et al.*, 1988; Scheidereit *et al.*, 1988; Staudt *et al.*, 1988; Sturm *et al.*, 1988). Homeobox sequences are also present in developmentally regulated mammalian genes (Dressler and Gruss, 1988; Holland and Hogan, 1988). The exact function and target sites of these genes during embryogenesis still remains to be determined, although it is assumed that their products are transcription regulatory proteins whose homeodomain interacts with specific DNA sequences (Dressler and Gruss, 1988; Thali *et al.*, 1988; Serfling, 1989). Since in mouse these genes are members of large families (Holland and Hogan, 1988), we investigated whether additional octamer-binding proteins could be identified. Moreover, we were particularly interested in finding octamer-binding proteins that are developmentally regulated.

The data presented in this report show that several additional proteins, binding specifically to the octamer motif, are present in various adult organs and in mouse embryos at different stages of development. Oct3 is detected in extracts of 12-day mouse embryos and in adult brain. Two maternally expressed octamer-binding proteins, Oct4 and Oct5, are also found during early embryogenesis. Both are present in unfertilized oocytes and embryonic stem cells, whereas only Oct4 is found in male and female primordial germ cells (PGC). Neither Oct4 nor Oct5 are detectable in sperm. Instead, a protein similar to the B-cell Oct2 is present. Therefore Oct4 expression is specific for the female germline at later stages of germ cell development. Oct2-like proteins are also found in embryo, brain and kidney. However, despite their identical electrophoretic mobility, they exhibit different binding properties.

Results

A family of octamer-binding proteins is present in nuclear extracts of different murine tissues

Nuclear extracts were prepared from 12-day mouse embryo, placenta, yolk sac and a variety of adult mouse tissues including lymphoid and non-lymphoid cells. The presence of octamer-binding proteins was tested with the electrophoretic mobility shift assay (EMSA) using a radioactively labelled fragment of the immunoglobulin heavy chain gene (IgH) enhancer containing the octamer motif (oligonucleotide o-wt; Figure 1). As described previously, the ubiquitous and the B-cell-specific proteins, referred to as Oct1 and Oct2 respectively, are identified (Singh *et al.*, 1986; Staudt *et al.*, 1986; Figure 2). Oct1 is present in all tissues, although amounts vary slightly. Oct2 is found in the human B-cell

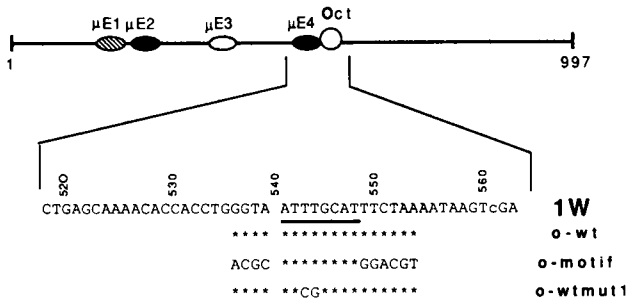

Fig. 1. A schematic representation of the mouse heavy chain gene enhancer (IgH) and the oligonucleotides that were used in the EMSA. The ellipses and the circle, designated μE1–μE4 and Oct respectively, represent binding sites identified by genomic sequencing (Church *et al.*, 1985; Ephrussi *et al.*, 1985). The original enhancer was isolated as a 997 bp *Xba*I–*Xba*I fragment (Banerji *et al.*, 1983). The 1W fragment containing the μE4- and Oct-binding sites was shown to be the smallest fragment conferring features of a B-cell-specific enhancer (Gerster *et al.*, 1987). The position of the 1W fragment and the positions of other oligonucleotides from the heavy chain gene enhancer are indicated in the figure. o-wt contains the octamer motif and surrounding sequences homologous to the IgH sequence. o-motif contains the octamer in the same position, but the rest of the oligonucleotide was replaced by random sequences. In o-wtmut1 two base changes were introduced in the octamer of o-wt. The numbering is according to Ephrussi *et al.* (1985). Asterisks indicate bases homologous to the IgH enhancer.

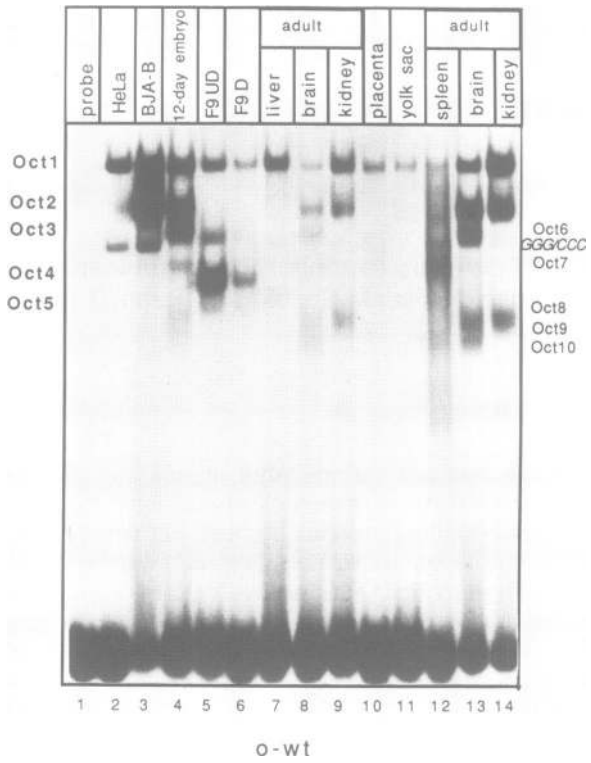

Fig. 2. Distribution of octamer-binding proteins in various nuclear extracts from adult mouse tissues, 12-day mouse embryos, lymphoid and non-lymphoid cell lines. Radiolabelled o-wt (**lane 1**) was incubated with nuclear extracts of the following origin. HeLa: human cervical carcinoma cell line (**lane 2**); BJA-B, human lymphoblastoid cell line (**lane 3**); 12-day embryos (**lane 4**); undifferentiated F9 (F9 UD): embryonic carcinoma cell line (**lane 5**); differentiated F9 (F9 D) (**lane 6**); liver (**lane 7**); brain (**lanes 8** and **13**); kidney (**lanes 9** and **14**); 12-day placenta (**lane 10**); 12-day yolk sac (**lane 11**); spleen (**lane 12**). For each lane 4 μg extract was used, except for lane 11 (1.4 μg), lane 12 (9 μg), lanes 13 and 14 (12 μg).

line BJA-B (lane 3 in Figure 2) and in the mouse myeloma cell line P3X63Ag8 (not shown). Surprisingly, an Oct2-like complex is detected in extracts of a 12-day-old embryo as well as in brain and kidney of adult mice.

Besides the two previously described proteins, a number of additional proteins are detected in the EMSA (Figure 2). These proteins were named Oct3–Oct10.

A strong binding activity, Oct3, differing from Oct1 and Oct2, was found in 12-day embryos and brains of adult mice (Figure 2, lanes 4, 8 and 13). Furthermore, in these extracts four faint complexes were found (Oct7–Oct10; lanes 4, 8, 13). The faint complexes become more obvious when higher amounts of protein are used (lane 13) or when a larger oligonucleotide is employed (Figures 4 and 5). Oct8 and Oct9 are also found in the kidney extract (lanes 9 and 14). In the extra-embryonic tissues – placenta and yolk sac – only Oct1 is detectable (lanes 10 and 11).

In spleen, many bands giving a long smear are observed with nuclear extracts (Figure 2, lane 12). Possibly a large number of proteins interact weakly with the octamer sequence. This is specific for the octamer motif and not due to the extract preparation since incubation of other oligonucleotides with spleen nuclear extracts gives rise to discrete bands (not shown). A complex running in front of Oct2 is detected in both human cell lines and, very weakly, in mouse liver (GGG–CCC complex, Figure 2, lanes 2, 3 and 7). This complex is due to a protein present in certain tissues recognizing either GGG or CCC protruding ends which were added to all oligonucleotides, except 1W.

Embryonic carcinoma (EC) cell lines are derived from the stem cells of teratocarcinomas and are a model system to study early events in mouse development. The stem cells of the EC line F9 can be differentiated into either parietal or visceral endoderm-like cells, depending on the culture conditions (for refs see Hogan *et al.*, 1986). Nuclear extracts from F9 stem cells and from differentiated parietal endoderm-like cells were analysed for octamer-binding proteins. Besides the ubiquitous Oct1 protein, three new complexes of distinct mobility are detected in F9 stem cells; the strongest one was named Oct4 (Figure 2, lanes 5 and 6). In front of Oct4, a faint complex, named Oct5, is detectable. This complex is easily observed when a longer fragment and higher amounts of extract are employed (see below). Another faint complex, called Oct6, runs at a position slightly ahead of that of the embryonal Oct3 complex. The difference in mobility is accentuated after prolonged gel electrophoresis (not shown). The intensities of all complexes decrease after differentiation of F9 stem cells. The decrease, however, is more dramatic for Oct4, Oct5 and Oct6 than for Oct1 (Figure 2, lanes 5 and 6).

The Oct proteins described above are not due to degradation of Oct1 and/or Oct2 since incubation of other oligonucleotides with the different nuclear extracts gives rise to single bands, indicating that there is no significant proteolysis during the extract preparation (not shown; see also Schöler *et al.*, 1989)

The patterns obtained with o-wt were reproduced with an oligonucleotide that also contains the μE4 binding site (1W; Figure 1). The patterns obtained with 1W and with o-wt were identical, except for the GGG–CCC complex missing in the 1W pattern (not shown). A complex that might correspond to μE4 was not detectable with this

oligonucleotide. 1W was used in the following experiments instead of o-wt, since in case of 1W the DNA−protein complexes are clearly separated, e.g. compare complexes Oct8−Oct10 in Figures 2 and 5.

The newly identified proteins interact with the octamer motif

Two additional oligonucleotides were synthesized, namely o-motif and o-wtmut1, in order to confirm that binding of the Oct proteins is due to the octamer motif (Figure 1). In o-motif, the octamer was kept in the same location, but the rest of the oligonucleotide was replaced by random sequences. As a result, the underlined T residue found in the decamer consensus sequence (TNATTTGCAT) is absent (Falkner and Zachau, 1984). This T residue found in front of many octamer sequences (Falkner *et al.*, 1986) is important for lymphoid-specific stimulation of transcription (Wirth *et al.*, 1987). In o-wtmut1, two base changes were introduced in the center of the octamer. With o-motif, the same qualitative pattern is obtained as with o-wt (Figure 3). The intensity of the Oct1 band is not affected, indicating that the octamer by itself is sufficient for the interaction. The binding of Oct2 and of the other proteins decreases, suggesting that flanking sequences affect binding. In case of B-cell Oct2, this decrease is about 2- to 3-fold, whereas all other proteins bind ~10-fold less. With o-wtmut1, however, only the GGG−CCC complex remains. In addition, a new band (complex N) with a slightly higher mobility than found for the Oct4 complex is visible in F9 and embryonal extracts (Figure 3, lanes 4−6 of o-wtmut1). Apparently, this protein interacts weakly with the mutant oligonucleotide. Complex N remains constant during differentiation of F9 cells.

The specificity of these proteins for the octamer motif was confirmed by competition experiments. In these experiments the oligonucleotides o-wt and o-motif, but not o-wtmut1, competed very efficiently against radiolabelled 1W. Corresponding to the binding results described above, 8-fold excess of o-wt competed similarly as a 40-fold excess of o-motif, whereas even a 200-fold excess of o-wtmut1 did not affect binding of the Oct-proteins (data not shown).

B-cell and brain Oct2 complexes differ in their stability at different temperatures

As indicated above, all Oct proteins, except Oct1, have different affinities for o-wt and o-motif (Figure 3). However, binding of B-cell Oct2 is less affected than binding of embryonal or brain Oct2 when the surrounding sequences of the octamer motif are changed (Figure 3, lanes 3 and 4 of o-wt and o-motif; brain not shown). To test if the Oct2 complexes can be further distinguished due to their binding requirements, BJA-B, embryonal and brain extracts were incubated at different temperatures. In this experiment, the extracts and the DNA were first warmed separately at 25, 37, 45, 55 and 65°C. After 5 min, DNA and extract were mixed and incubated an additional 15 min. The BJA-B Oct2 complex clearly shows optimum binding at 37°C (Figure 4). The brain (Figure 4) and the embryonal Oct2 (not shown) complexes, however, remain constant between 25 and 45°C. In each case binding of Oct1 is not affected in this temperature range. At temperatures >45°C the Oct proteins are heat inactivated in each extract. To test if the decreased binding of the BJA-B Oct2 is due to the stability of the

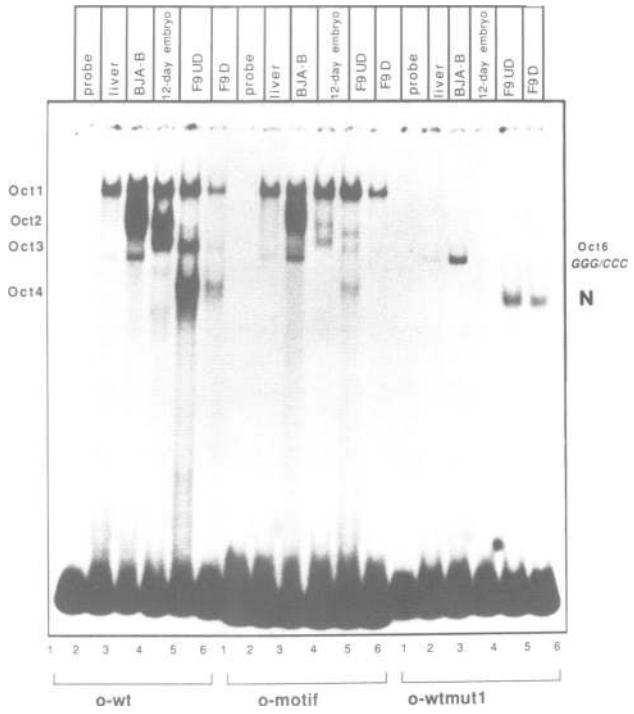

Fig. 3. Specificity of binding to the octamer motif. The radiolabelled oligonucleotides o-wt (left), o-motif (centre) or o-wtmut1 (right) were incubated in the absence (**lanes 1**) or in the presence of nuclear extracts of the following origin: liver (**lanes 2**); BJA-B (**lanes 3**); 12-day embryos (**lanes 4**); F9 (UD) (**lanes 5**); F9 (D) (**lanes 6**). For each lane 4 μg extract was used.

DNA−protein complex at this temperature or to heat inactivation or degradation of the protein, the BJA-B extract was first heated at 45°C and then further incubated at 25, 37 and 45°C. The same pattern with almost the same binding efficiencies was obtained in this experiment and in the one described above, indicating that the BJA-B Oct2 is not heat inactivated or degraded at 45°C. Thus, besides their affinity to different oligonucleotides, both Oct2 complexes are distinguishable by their stability at different temperatures.

Octamer-binding proteins are unequally distributed in brain

To test whether Oct proteins are region-specifically expressed in the brain, the procedure for extract preparation was altered to accommodate small cell samples (see Materials and methods). With this microextraction procedure, whole-cell extracts of several rat brain regions were prepared. Rat brain was used since it is dissected easily into brain subregions without any cross-contamination and because the Oct patterns of mouse and rat brain are identical (not shown). Similar patterns are observed in the EMSA when total brain is compared with subregions (Figure 5, left). However, the subregions show quantitative and also qualitative differences when compared to each other. A quantitative difference is most obvious for the cerebellum where only minor amounts of Oct2, Oct3 and Oct7 are detectable. In the frontal cortex and in the subcortical region, Oct3 is accompanied by an additional band of slightly higher mobility.

When extracts of 12-day embryonal heads and bodies are prepared with the microextraction procedure, higher amounts of Oct2, Oct3 and Oct7 are found in the head region than in the rest of the body (Figure 5, right). At this stage it is

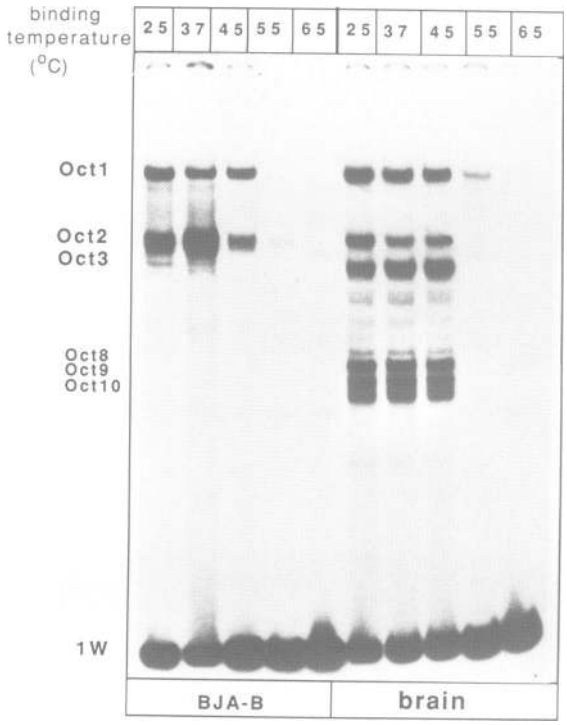

Fig. 4. Stability of brain and B-cell Oct complexes at different temperatures. Radiolabelled 1W was incubated at different temperatures with BJA-B (left) or brain (right) nuclear extracts. The temperatures are indicated at the top of the figure. For each lane 4 μg of nuclear extract was used.

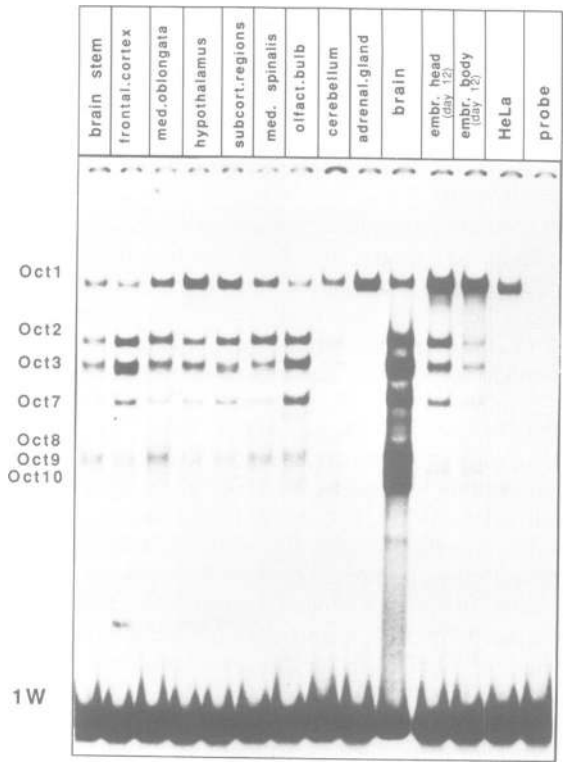

Fig. 5. Distribution of octamer-binding proteins in rat brain and in 12-day embryo. Radiolabelled 1W (right) was incubated with whole-cell extracts of different rat brain regions, 12-day embryonal head or 12-day embryonal body as indicated at the top of the figure. Nuclear extracts of HeLa cells and mouse brains are included. For each lane 1 μg extract was used, except for HeLa (0.5 μg).

not possible to restrict these proteins to certain parts of the embryonal body, although they are absent in certain tissues (see below).

Oct4 and Oct1 have structurally related DNA-binding domains

To test the relationship of Oct4 with Oct1 and Oct2, proteolytic clipping experiments were performed (Schreiber *et al.*, 1988). Nuclear extracts of HeLa, F9 and BJA-B cells were first incubated with 1W and subsequently digested with different amounts of dispase or trypsin. Dispase was chosen because it is a non-specific endoprotease and should detect structural similarities between the Oct proteins. Trypsin is a specific protease that cleaves after arginine or lysine residues. Therefore, this protease should emphasize, despite possible structural similarities, differences in their amino acid sequence.

Dispase digestion of each extract resulted in small fragments of the same size which retained DNA binding activity (Figure 6A). A comparison of the proteolytic intermediates reveals that the patterns found with the HeLa and F9 extracts are more similar to each other than to the BJA-B pattern. This similarity suggests a structural relationship between Oct1 and Oct4.

The data obtained with dispase were complemented with trypsin clipping experiments. Again, small fragments retaining DNA binding activity were generated (Figure 6B). However, trypsin digestion reveals differences between these fragments. With the BJA-B extract two fragments are obtained (core 2a and core 2b), whereas with both other extracts, only one fragment shows DNA-binding activity (core 1 and core 4). Core 2a and core 2b decrease gradually when trypsin amounts >0.3 ng are used, with core 2a remaining the strongest band. Core 1 and core 4 have the same mobility, running between core 2a and core 2b.

Similar to the digestion by dispase, Oct4 is degraded in multiple steps. In this case, however, only Oct1 has two intermediates. The faint band in the HeLa extract is not an intermediate but an undefined product remaining in all lanes and is also found in the presence of dispase. Digestion of Oct2-containing extracts yields smaller intermediates than with the other extracts. Therefore the proteolytic intermediates are different for each extract when trypsin is used. These experiments also show that apart from their DNA-binding domain, Oct1 and Oct4 have structural differences.

Oct4 is present in the female germ cell lineage

Although the F9 cell system has been extensively used, these EC cells have been subjected to possibly deleterious selection pressures. They are restricted in their developmental capacities and their exact stage of differentiation is unknown. Therefore we extended our experiments to pluripotent embryo-derived stem (ES) cells. ES cells are similar to EC cells in several aspects (Evans and Kaufman, 1981; Martin, 1981). However, they have been directly isolated from the inner cell mass (ICM) of blastocysts and they remain pluripotent, contributing to all somatic and germ cell lineages when returned to the embryonic environment (Williams *et al.*, 1988a and references therein). When nuclear extracts of D3 (Doetschman *et al.*, 1985) and F9 cells were compared, the shifting pattern was identical, indicating that the blastocyst-derived ES cell line D3 also contain Oct4, Oct5 and Oct6 (not shown).

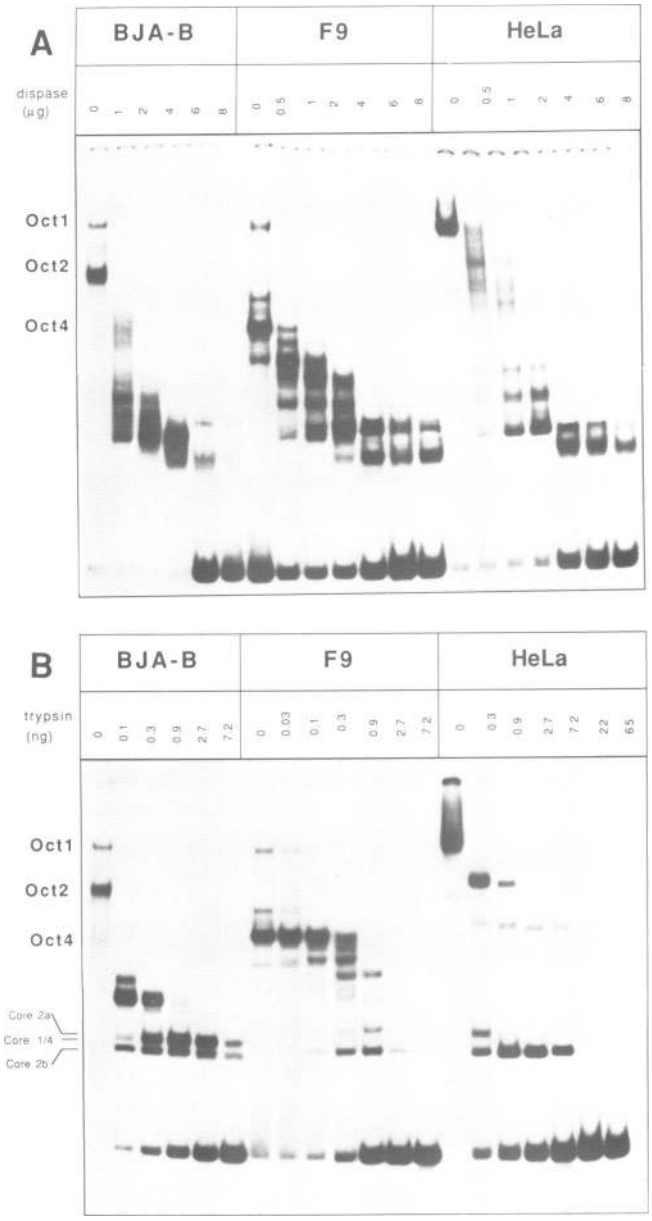

Fig. 6. Proteolytic clipping EMSA of octamer DNA−protein complexes. Radiolabelled 1W was incubated first with BJA-B (left), F9 (centre) or HeLa (right) nuclear extracts. After 15 min different amounts of dispase (**A**) or trypsin (**B**) were included. No protease was used for the first lane of each extract set. Oct1, Oct2 and Oct4 indicate the undegraded complexes. Core 1 and core 4 indicate the positions of the Oct1 and Oct4 DNA-binding domains respectively. Core 2a and core 2b show the positions of the DNA binding domains after Oct2 clipping. For each lane 6 μg nuclear extract was used.

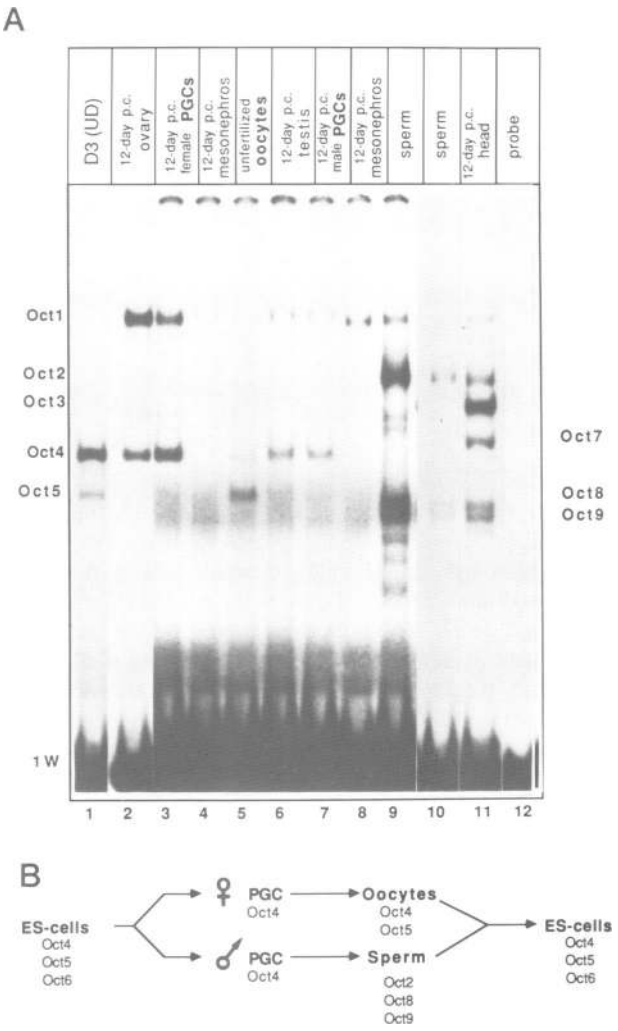

Fig. 7. Germline-specific expression of Oct proteins. (**A**) Radiolabelled 1W (**lane 12**) was incubated with whole-cell extracts of the following origin: undifferentiated D3 cells (D3 UD) (**lane 1**), 12-day female embryo genital ridges (**lane 2**), PGCs of 12-day female embryo genital ridges (**lane 3**), 12-day embryonal mesonephros (**lanes 4** and **8**), unfertilized ovulated oocytes (**lane 5**), 12-day male embryo genital ridges (**lane 6**), PGCs of 12-day male embryo genital ridges (**lane 7**), mature sperm (**lanes 9** and **10**) and 12-day embryonal head (**lane 11**). Lanes 3−9 show 48-h exposures, the rest 16-h exposures.
(**B**) Summary of the distribution pattern of octamer-binding proteins as it is deduced from panel (A). Oct1 is not included.

Since Oct4, Oct5 and Oct6 are already present in early embryogenesis, stages preceding fertilization were investigated for Oct proteins. Whole-cell extracts were prepared from ~1000 unfertilized oocytes and from mature sperm with the microextraction procedure and compared with a D3 whole-cell extract. The patterns of D3 cells and oocytes are similar, i.e. oocytes contain Oct1, Oct4 and Oct5 (Figure 7A, lanes 1 and 5). However, Oct5 forms the strongest band in the female gametes. Therefore, the intensities of the Oct4 and the Oct5 complexes are inverse in ES cells and oocytes. These proteins were not detectable in sperm, where Oct2, Oct8 and Oct9 give the most prominent complexes (Figure 7A, lanes 9 and 10).

Both gametes are derived from primordial germ cells (PGCs) that differentiate from cells of the primitive ectoderm (Copp *et al.*, 1986). The PGCs are first identified in the 8-day embryo at the base of the allantois. From there they migrate to and colonize the genital ridge of the presumptive gonad, which is situated near the kidney. Their numbers increase from ~10 to ~100 at day 8 to ~25 000 PGCs at day 13 when the genital ridges are fully colonized (Tam and Snow, 1981). To test these cells for Oct proteins, PGCs were obtained from male and female genital ridges of 12-day embryos (Hogan *et al.*, 1986). The genital ridges were dissected from the embryo together with the mesonephros that lies on the gonads like a shield. Subsequently both organs were separated. About 1000 PGCs from ovaries and ~400 PGCs from testes were collected and extracted with the

Table I. Distribution of octamer-binding proteins in the mouse

	Liver	Brain	Adrenal gland	Kidney	Embryonal head	Embryonal body	Yolk sac	Placenta	Meso-nephros	PGC ♀	PGC ♂	Sperm	Oocytes	D3	F9	Ag8
Oct1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Oct2		+		+	+	+						+				+
Oct3		+			+	+										
Oct4										+	+		+	+	+	
Oct5													+	+	+	
Oct6														+	+	
Oct7		+			+	+										
Oct8		+		+	+	+						+				
Oct9		+		+	+	+						+				
Oct10		+			+	+										

microextraction procedure. In parallel, the mesonephric tissue and the remaining genital ridges were also extracted.

Both male and female PGCs contain Oct4 as the most abundant octamer-binding factor in these cells (Figure 7A, lanes 3 and 7). Although Oct5 is found in later stages of development, it is not detectable in PGCs. The patterns of the PGCs and their respective genital ridges are identical, with higher amounts of Oct1 in the remaining organs (Figure 7A, lanes 2, 3 and lanes 6, 7). The presence of Oct4 is probably due to PGCs left in the genital ridges, although the presence of Oct4 in somatic cells of testis and ovary cannot be excluded. In the mesonephric tissue, only Oct1 is detectable (Figure 7A, lanes 4 and 8).

Since Oct4 and Oct5 are found in oocytes but not in sperm, both would be the first maternally expressed transcription factors described so far (see also Schöler *et al.*, 1989). Moreover, the presence of Oct4 in PGCs and unfertilized oocytes suggest that this protein is specifically expressed in cells representing the germline. Male and female gametes are distinguishable by the presence of Oct2 or Oct4 respectively (Figure 7B).

Discussion

A family of octamer-binding proteins has been identified in various mouse tissues and in embryo at different stages of development. In addition to the previously described Oct1 and Oct2, we show that eight other proteins, Oct3–Oct10, are found in certain tissues of the mouse (Table I). The specificity of all factors was shown using mutated binding sites and competition experiments. It is not known, however, if these complexes are due to the products of distinct genes, or if some of these represent post-translational modifications of either Oct1 and/or Oct2. Alternatively, differential splicing could give rise to distinct products interacting with the octamer motif.

Oct2 is thought to be restricted to B-cells (Gerster *et al.*, 1987 and refs. therein). Surprisingly, Oct2-like complexes are also detected in extracts of embryo, sperm, brain and kidney. However, as demonstrated for the brain and embryonic Oct2 complexes, they are distinguishable from B-cell Oct2 by their stability at different temperatures and by their affinity to different oligonucleotides. These experiments cannot distinguish if the respective binding properties are due to a modification of one Oct2 protein or to different proteins. Furthermore, the experiments show that the Oct1 and Oct2 proteins in B-cells have differential binding requirements.

Additional support for the presence of distinct Oct2-like proteins was obtained by trypsin-clipping experiments. B-cell extracts yielded two distinct DNA-binding domains, named core 2a and core 2b. Again, it cannot be excluded that both bands are derived from one protein. However, the core 2a–DNA and core 2b–DNA complexes decrease simultaneously when increasing amounts of trypsin are used, suggesting that both are formed independently. Future experiments will be directed towards resolving if core 2a and core 2b are due to one or two B-cell proteins and also if one (or both) of them is related to the Oct2-like factor(s) found in other tissues.

Recently, OTF-2B has been described as a second B-cell-specific protein (Schreiber *et al.*, 1988). The authors demonstrated that in the EMSA OTF-2B runs as a faint complex between the ubiquituous and the other B-cell-specific factor. At a similar position we also detected a faint band, most probably indicating OTF-2B (e.g. in Figure 6). Since the trypsin digestion of our B-cell extract gives rise to well-detectable DNA-binding domains, it is unlikely that core 2a or core 2b is derived from the OTF-2B complex.

Oct3 is found in the adult brain and in the embryo. In our experiments, Oct3 is always accompanied by Oct2 and Oct7. Quantitative and qualitative differences of Oct1, Oct2, Oct3 and Oct7 complexes could be demonstrated in different brain subregions. For this comparison, equal amounts of extract were used not taking into account immense size differences of neural cell types. Despite these size differences, however, the amounts of the ubiquitously present Oct1 protein varied only slightly in our experiments. Strikingly, Oct2, Oct3 and Oct7 are much less abundant in the cerebellum when compared to the other regions. However, since each of the subregions is composed of different cell types, a defined correlation of these Oct factors with certain neuronal cells is not possible. The cerebellum, for example, has five different neuronal types. To identify the cell type(s) expressing these factors, a immunocytochemical analysis of brain sections using specific antibodies to these factors or established brain cell lines will be necessary. Because the Oct patterns identified in mouse and rat brain extracts are similar, both tissues can be used as a source. In addition to the quantitative differences, qualitative differences are found in the brain subregions. In the frontal cortex and in the subcortical regions an additional band is detected running

slightly ahead of the Oct3 complex. This position is similar to that of the Oct6 complex found in F9 cells, although a difference in migration is obvious after a prolonged gel electrophoresis.

In the embryonal head region, Oct2, Oct3 and Oct7 are more abundant than in the body. Since these proteins are found in the adult brain, the central nervous system is also a prime candidate for their presence in the embryo. In contrast to the various proteins found in the embryo, in the extra-embryonic tissues—placenta and yolk sac—only Oct1 is found.

Early in mouse development, certain cells are set aside as progenitors of the germ cells. These primordial germ cells migrate to the developing gonads, where after a period of mitotic proliferation, they undergo meiosis and differentiation into oocytes and sperm. Oct4 is found in embryonic stem cells, in primordial germ cells and in unfertilized oocytes, and thus is present in cells that ensure the continuous flow of genetic information. The presence of Oct4 not only in germ cells but also in embryonic stem cells suggests that this protein is present in cells having the potency to produce germ cells. In this respect, Oct4 might turn out to be a marker of the germ cell lineage.

In contrast to Oct4, Oct5 is only detectable in oocytes and embryonic stem cells. Their presence in oocytes and in ES cells suggests that Oct4 and Oct5 are not only present as maternal proteins in the oocyte but can also be synthesized *de novo* by zygotic expression in the embryo. Since Oct5 reveals the most prominent complex in oocytes and is barely detectable in ES cells, both developmental stages seem to have opposite amounts of Oct4 and Oct5. Therefore it will be interesting to investigate if the different ratios of these proteins reflect regulatory events during early mouse embryogenesis. Although the biological significance has to be determined, it is striking that Oct4 is present in cells having embryonic stem cell properties such as PGCs, oocytes, ES and EC cells.

The presence of Oct4 and Oct5 in F9 stem cells and their decrease during differentiation renders this cell line an ideal model system for the functional analysis of these proteins. Recently, Lenardo *et al.* (1989), have also found a factor named NF-A3 which is present in F9 cells and differs from Oct1. Although we detect two more complexes in F9 extracts, NF-A3 is most probably equivalent to Oct4 since both proteins give the most prominent complexes. However, their functional analysis of the octamer motif in F9 cells results in different conclusions than our analysis (Schöler *et al.*, 1989).

In sea urchin a testis-specific octamer complex has been described (Barberis *et al.*, 1987). In addition, an octamer-binding protein has been identified in malignant melanoma cells (Cox *et al.*, 1988). Since these experiments were done independently, it is not possible at this point to attribute these proteins to any of the complexes described here.

The transcription factors Oct1 and Oct2 both contain a region homologous to the homeodomain that is present in a variety of other vertebrate genes. Oct3, Oct4 and Oct5 bind to the same motif and their binding is also abolished when the octamer motif is mutated, suggesting that they might also have a homeodomain. Proteolytic clipping experiments with F9, HeLa and BJA-B extracts demonstrate that Oct1 and Oct4 have structurally very similar DNA binding domains. Cloning of the Oct4 cDNA will show if this similarity can be extended to sequence homology. Cloning of the respective gene(s) may also reveal if distinct genes or differential splicing give rise to these different Oct proteins and may allow a functional analysis of these proteins.

Materials and methods

Cell lines and transfection

F9, BJA-B and HeLa cells were grown as described previously (Schöler and Gruss, 1984, 1985). D3 cells were cultured by propagating them in highly purified, recombinant LIF to maintain the undifferentiated phenotype (Williams *et al.*, 1988b). The F9 (UD) cells were differentiated to F9 (D) by treating them 48 h before harvest with 5×10^{-7} M retinoic acid and 10^{-3} M cAMP (Strickland *et al.*, 1980).

Isolation of 12-day primordial germ cells

The isolation of PGCs was carried out essentially as described by Hogan *et al.* (1986): genital ridges were dissected out from embryos around day 12.5 when the male and female genital ridges can be distinguished (Hogan *et al.*, 1986), separated from the mesonephric shield and incubated in an EDTA–saline solution at room temperature for 15 min. Developing ovaries and testes were incubated separately to allow isolation of female and male PGCs. The PGCs were released from the developing gonads after EDTA treatment by puncturing the genital ridges with a hypodermic needle, collected and immediately frozen in liquid nitrogen.

Isolation of ovulated, unfertilized oocytes

Six-week-old NMRI female mice were superovulated by injecting i.p. 5 units of gonadotropin from pregnant mare serum 48 h prior to injecting 5 units of human chorionic gonadotropin (HCG). Eighteen hours after HCG injection, ovulated oocytes were flushed from the oviducts with M2 medium (Hogan *et al.*, 1986), freed of cumulus cells by hyaluronidase treatment and frozen in liquid nitrogen.

Isolation of sperm

The epididymis from male C57Bl/6 mice were dissected out and placed into PBS. Sperm was gently released from the epididymis and allowed to swim freely for ~15 min. Epididymal tissue was separated from sperm by centrifugation at 500 *g* for 5 min. Sperm was pelleted by centrifugation at 2000 *g* for 5 min and frozen in liquid nitrogen.

Preparation of nuclear extracts

Nuclear extracts of the cell lines were prepared according to the protocol of Dignam *et al.* (1983). Nuclear extracts from NMRI mouse tissues and embryos were prepared according to the method of Gorski *et al.* (1986) including a mixture of protease inhibitors: PMSF (0.5 mM), leupeptin (0.5 μg/ml), pepstatin (0.7 μg/ml), aprotinin (1 μg/ml) and bestatin (40 μg/ml) in the homogenization buffer.

Microextraction procedure for the preparation of whole cell extracts

As little as 400 cells were prepared by dissolving the cellular pellet or piece of tissue in 100 μl extraction buffer. The extraction buffer contained 20 mM Hepes, pH 7.8 (4°C), 450 mM NaCl, 0.2 mM EDTA, 0.5 mM DTT, 25% glycerol and the same protease inhibitors as listed above. After sonication with a Branson Cell Disruptor (small tip; 10 pulses; output control 3; 30% duty cycle) the extracts were cleared by centrifugation in an Eppendorf centrifuge. Up to 8 μl of the extract were used directly in the EMSA. All steps were performed on ice or in the cold room. In case of preparing different regions of the brain, the time from sacrificing the animal to loading the gel for the EMSA took ~45 min.

The procedure was set up with the octamer motif using several conditions and different amounts of BJA-B cells. The final procedure is advantageous for several reasons: it is very fast, only small quantities of protease inhibitors are needed and the inhibitors can be included from the beginning of the preparation. Most importantly, small samples of <1000 cells can be extracted to yield a whole-cell extract. To test the reliability of the extraction procedure we examined the stability of the Oct proteins in these extracts. Since the brain nuclear extract showed most bands in the EMSA, this tissue was tested for potential degradation. After 1 h of pre-incubation at 37°C without DNA the pattern of a brain extract did not change (not shown). Therefore the new procedure yields a product that is at least stable enough to perform the EMSA. Only quantitative differences can be detected between the nuclear and the whole-cell extract. In particular, Oct8, Oct9 and Oct10

are stronger in the nuclear extract but also detectable in the whole-cell extract (see Figure 5 where only the mouse brain represents a nuclear extract). From this we conclude that the results obtained with the quick microextraction procedure are reliable.

The protein concentration was determined as described by Kalb and Bernlohr (1977).

Electrophoretic mobility shift assay (EMSA) and proteolytic clipping assay

Binding conditions for the EMSA were 10 mM Hepes, pH 7.8 (4°C), 1 mM spermidine, 5 mM $MgCl_2$, 50 mM KCl, 0.5 mM DTT, 9% glycerol, 1 μg poly d(I·C), 10 000 c.p.m. of the labelled oligonucleotide; the final reaction volume was 15 μl. After an incubation for 15 min at 25°C the binding reaction was applied to gel electrophoresis as described by Singh *et al.* (1986).

In the protease-clipping assay binding is performed as in the EMSA. Different amounts of proteases are then added together with blue dye to the binding assays. After 5 min the samples are applied to the gel.

Oligonucleotides

The numbering of all oligonucleotides is according to Ephrussi *et al.* (1985). 1W is identical to region 518–564 except for 562 (Gerster *et al.*, 1987), o-wt is identical to the octamer motif containing region (536–554), o-motif is identical to the octamer motif (541–548) surrounded by random non-homologous sequences and o-wtmut1 is identical to o-wt except for base-pair exchanges at positions 543 and 544). Sequences homologous to the octamer motif are underlined. The sequences of the coding strands of 1W, o-wt, o-motif and o-wtmut1 are shown in Figure 1. Three additional G residues at the 5′ end of the coding strand and three additional C residues at the 5′ end of the noncoding strand were added to all oligonucleotides except for 1W.

Acknowledgements

We thank C.Peiker for excellent technical assistance, N.Gough for LIF, A.Zimmer for providing us with D3 cells, M.Olowson for help during the oocyte isolation, G.I.Kristjansson for dissecting the rat brain, A.Jacob for an additional F9 extract and R.Altschäffel for doing the photographic work. Furthermore we would like to thank A.McLaren for encouraging us to prepare the PGC extracts. We are grateful to A.Püschel, D.Siekhaus, H.Nornes and G.Dressler for valuable comments on the manuscript. A.H. was supported by EMBO and Alexander von Humboldt Fellowships. N.S. is a recipient of the Behrens-Weise Foundation. This project was supported by the Max-Planck Society and by the Bundesministerium für Forschung und Technologie (BMFT).

References

Banerji,J., Olson,L. and Schaffner,W. (1983) *Cell*, **33**, 729–740.
Barberis,A., Superti-Furga,G. and Busslinger,M. (1987) *Cell*, **50**, 347–359.
Church,G.M., Ephrussi,A., Gilbert,W. and Tonegawa,S. (1985) *Nature*, **313**, 798–801.
Clerc,R.G., Corcoran,L.M., LeBowitz,J.H., Baltimore,D. and Sharp,P.A. (1988) *Genes Dev.*, **2**, 1570–1581.
Copp,A.J., Roberts,H.M. and Polani,P.E. (1986) *J. Embyrol. Exp. Morphol.*, **95**, 95–115.
Cox,P.M., Temperley,S.M., Kumar,H. and Goding,C.R. (1988) *Nucleic Acids Res.*, **16**, 11047–11056.
Dignam,J.D., Lebowitz,R.M. and Roeder,R.G. (1983) *Nucleic Acids Res.*, **11**, 1475–1489.
Doetschman,T.C., Eistetter,H., Katz,M., Schmidt,W. and Kemler,R. (1985) *J. Embryol. Exp. Morphol.*, **87**, 27–45.
Dressler,G.R. and Gruss,P. (1988) *Trends Genet.*, **4**, 214–219.
Dreyfus,M., Doyen,N. and Rougeon,F. (1987) *EMBO J.*, **6**, 1685–1690.
Ephrussi,A., Church,G.M., Tonegawa,S. and Gilbert,W. (1985) *Science*, **227**, 134–140.
Evans,M.J. and Kaufman,M.H. (1981) *Nature*, **292**, 154–156.
Falkner,F.G. and Zachau,H.G. (1984) *Nature*, **310**, 71–74.
Falkner,F.G., Mocinat,R. and Zachau,H.G. (1986) *Nucleic Acids Res.*, **14**, 8819–8827.
Fletcher,C., Heintz,N. and Roeder,R.G. (1987) *Cell*, **51**, 773–781.
Gerster,T., Matthias,P., Thali,M., Jiricny,J. and Schaffner,W. (1987) *EMBO J.*, **6**, 1323–1330.
Gorski,K., Carneiro,M. and Schibler,U. (1986) *Cell*, **47**, 767–776.
Hatzopoulos,A.K., Schlokat,U. and Gruss,P. (1988) In Hames,B.D. and Glover,D.M. (eds), *Transcription and Splicing, Frontiers in Molecular Biology*. IRL Press, Oxford, pp. 43–96.
Hogan,B., Costantini,F. and Lacy,E. (1986) *Manipulating the Mouse Embryo*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Holland,P.W.H. and Hogan,B.L.M. (1988) *Genes Dev.*, **2**, 773–782.
Kalb,V.R. and Bernlohr,R.M. (1977) *Anal. Biochem.*, **82**, 362–371.
Ko,H.-S., Fast,P., McBride,W. and Staudt,L.M. (1988) *Cell*, **55**, 135–144.
Landolfi,N.F., Capra,J.D. and Tucker,P.W. (1986) *Nature*, **323**, 548–557.
Lenardo,M.J., Staudt,L., Robbins,P., Kuang,A., Mulligan,R.C. and Baltimore,D. (1989) *Science*, **243**, 544–546.
Martin,G.R. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 7634–7638.
Müller,M.M., Ruppert,S., Schaffner,W. and Matthias,P. (1988) *Nature*, **336**, 544–551.
Rosales,R., Vigneron,M., Macchi,M., Davidson,I., Xiao,J.H. and Chambon,P. (1987) *EMBO J.*, **6**, 3015–3025.
Scheidereit,C., Heguy,A. and Roeder,R.G. (1987) *Cell*, **51**, 783–793.
Scheidereit,C., Cromlish,J.A., Gerster,T., Kawakami,K., Balmaceda, C.-G., Currie,R.A. and Roeder,R.G. (1988) *Nature*, **336**, 551–557.
Schöler,H.R. and Gruss,P. (1984) *Cell*, **36**, 403–411.
Schöler,H.R. and Gruss,P. (1985) *EMBO J.*, **4**, 3005–3013.
Schöler,H.R., Hatzopoulos,A.K. and Schlokat,U. (1988) In Kahl,G. (ed.), *Architecture of Eukaryotic Genes*. Verlag Chemie, Weinheim, pp. 89–120.
Schöler,H.R., Balling,R., Hatzopoulos,A.K., Suzuki,N. and Gruss,P. (1989) *EMBO J.*, **8**, 2551–2557.
Schreiber,E., Matthias,P., Müller,M.M. and Schaffner,W. (1988) *EMBO J.*, **7**, 4221–4229.
Serfling,E. (1989) *Trends Genet.*, **5**, 131–133.
Singh,H., Sen,R., Baltimore,D. and Sharp,P.A. (1986) *Nature*, **319**, 154–158.
Staudt,L.M., Singh,H., Sen,R., Wirth,T., Sharp,P.A. and Baltimore,D. (1986) *Nature*, **323**, 640–643.
Staudt,L.M., Clerc,R.G., Singh,H., LeBowitz,J.H., Sharp,P.A. and Baltimore,D. (1988) *Science*, **241**, 577–580.
Strickland,S., Smith,K.K. and Marotti,K.R. (1980) *Cell*, **21**, 347–355.
Sturm,R.A., Das,G. and Herr,W. (1988) *Genes Dev.*, **2**, 1582–1599.
Tam,P.P.L. and Snow,M.H.L. (1981) *J. Embryol. Exp. Morphol.*, **64**, 133–147.
Tanaka,M., Grossniklaus,U., Herr,W. and Hernandez,N. (1988) *Genes Dev.*, **2**, 1764–1778.
Thali,M., Müller,M.M., DeLorenzo,M., Matthias,P. and Bienz,M. (1988) *Nature*, **336**, 598–601.
Williams,R.L., Courtneidge,S.A. and Wagner,E.F. (1988a) *Cell*, **52**, 121–131.
Williams,R.L., Hilton,D.J., Pease,S., Willson,T.A., Stewart,C.L., Gearing,D.P., Wagner,E.F., Metcalf,D., Nicola,N.A. and Gough,N.M. (1988b) *Nature*, **336**, 684–687.
Wirth,T., Staudt,L. and Baltimore,D. (1987) *Nature*, **329**, 174–178.

Received on May 5, 1989; revised on June 9, 1989