Nucleic Acids Research, Vol. 20, No. 21 5519–5525

Cloning by recognition site screening of two novel GT box binding proteins: a family of Sp1 related genes

Gustav Hagen, Susanne Müller, Miguel Beato and Guntram Suske*
Institut für Molekularbiologie und Tumorforschung, Philipps-Universität Marburg, Emil-Mannkopff-Straße 2, W-3550 Marburg, Germany

Received October 6, 1992; Accepted October 15, 1992 EMBL accession nos X68560, X68561

ABSTRACT

Previous analyses of the uteroglobin gene promoter revealed a GT1 box which is also found in the SV40 enhancer. The GT1 element in the context of the uteroglobin promoter is active in Ishikawa cells, a human endometrial cell line, but not in HeLa cells. Here we report the cloning by recognition site screening of two factors (SPR-1 and SPR-2) which bind to this GT1 motif. SPR-1 and SPR-2 are homologues of the transcription factor Sp1. All three proteins are closely related members of a gene family encoding proteins with very similar structural features. Like Sp1, SPR-1 and SPR-2 contain glutamine and serine/threonine rich amino acid stretches. Most significantly, the DNA binding domains of all three proteins are highly conserved and they recognize GT as well as GC boxes identically. SPR-2 mRNA is expressed ubiquitously, whereas SPR-1 transcripts are abundant in the brain but barely detectable in other organs. The possible function of these factors for the activity of the uteroglobin promoter is discussed.

INTRODUCTION

The combinatorial action of transcription factors which bind to promoter and enhancer elements specifies the expression pattern of a gene in the organism. In the last few years, we have begun to explore the mechanisms responsible for the expression of the rabbit uteroglobin gene in certain ontogenetically unrelated epithelial cell types (mainly in endometrium and lung). We found that the uteroglobin promoter is utilized 10-fold more efficiently in the human endometrial cell line Ishikawa and in the lung cell line NCI-H441 than in HeLa cells or fibroblasts (1,2). Gene transfer experiments with 5′-promoter deletions and linker scanning promoter mutations as well as *in vitro* DNA binding studies with nuclear extracts identified several *cis* acting elements that are involved in promoter activity in Ishikawa cells. At least seven mutation-sensitive regions of the uteroglobin promoter contribute to its overall activity in this endometrial cell line (1,2) (Fig. 1). One of these regions located 230 bp upstream of the transcription start site contains a sequence which is in 12 out of 13 nucleotides identical to the GT1 motif of the SV40 enhancer (3,4,5). It also contains a so called CACCC box found in several other enhancers and promoters, including the tyrosine aminotransferase gene (6,7), the tryptophan oxygenase gene (8,9) and the β-globin gene (10). The GT1/CACCC element in the context of the uteroglobin promoter is active in Ishikawa cells but not in HeLa cells (1). Although some CACCC boxes are known to be recognized by Sp1 (11), to date other factors that bind to these elements are not well characterized or cloned. To understand the molecular basis for the cell specificity of this element in the context of the uteroglobin promoter we have decided to clone the cDNA(s) of the protein(s) binding to the GT1 motif.

Here we report the cloning by recognition site screening of two novel factors which bind to the GT1 motif of the rabbit uteroglobin promoter. Both proteins are homologues of the transcription factor Sp1. Therefore, we designate them Sp1-related factors 1 and 2 (SPR-1 and SPR-2). All three proteins have similar structural domains and recognize GT and GC boxes with identical affinities. Northern blot analyses reveal that SPR-2 mRNA is expressed ubiquitously like Sp1 mRNA. In contrast, SPR-1 transcripts are abundant in brain but barely detectable in other organs.

MATERIALS AND METHODS

Construction of cDNA expression libraries and screening procedure

Two cDNA expression libraries in λ gt11 were constructed essentially as described (12,13) using poly(A)-enriched RNA from Ishikawa cells (14) and commercially available random and oligo (dT) primers. For insertion of the cDNA into the EcoRI cloning site of λ gt11 DNA a synthetic adapter containing a NotI site adjacent to an EcoRI site was used (15). The random and oligo(dT) primed libraries contain 10^6 and 10^7 independent p.f.u., respectively. Ninety percent of the phages contain recombinant λgt11 DNA. Initially, the random primed library was screened according to Singh *et al.* (16) with the modifications of Vinson *et al.* (17) using concatemers of the ^{32}P-labeled double stranded oligonucleotide which has the sequence 5′-ACCCCTTGCCACACCCCTGCACAAG-3′.

* To whom correspondence should be addressed

Plasmid constructions

All λgt11 phage cDNA inserts (see Fig. 2 A and C) were subcloned into the NotI site of the plasmid vector pBluescript II KS− (pKS) and sequenced by the chain termination method (18) using denatured double stranded DNA templates (19). To express zinc finger containing fragments of SPR-1, SPR-2 and Sp1 in *E.coli* under the control of the T7 RNA polymerase dependent φ10 promoter (20), expression plasmids were constructed as follows. The plasmids for the expression of SPR-1 and SPR-2 were generated by inserting a 3 kb SmaI fragment from plasmid pKS-A8O and a 0.8 kb blunted AccI-EcoRI fragment from pKS-A3O into the blunted BamHI sites of pET-3b and pET-3c, respectively. The Sp1 expression plasmid was constructed by cloning a blunted 4.4 kb XbaI fragment from the plasmid pBS-Sp1-fl (kindly provided by Dr. R. Tjian) into the bluntrd BamHI site of pET-3c.

Ligation mediated PCR

The conditions for the synthesis of double stranded cDNA were essentially as described (12) using total RNA from Ishikawa cells (20 μg) and a specific primer (5′-GAGTCCCTATTTTGCTGC-AAGTAGCTGCCAG-3′) corresponding to nucleotides 332 to 362 of the SPR-1 cDNA shown in Fig. 2 B). The cDNA synthesis products were purified by Sephadex G-50 chromatography and concentrated by lyophilization. A synthetic linker obtained by hybridization of the oligonucleotides 5′-GCGGTGACCCGGG-AGATCTGAATTC-3′ (LMPCR1) and 5′-GAATTCAGATC-3′ (LMPCR2) was ligated to the cDNA. For the amplification reaction the oligonucleotides 5′-GCCAGAGGAGAGGGCTG-AGAGTCCT-3′ (corresponding to nucleotides 299 to 327 in Fig. 2 B) and LMPCR1 (see above) were used as primers. The ligation of the linker and the amplification of the cDNAs have been performed under conditions essentially as described (21,22). The PCR products were digested with BglII (cuts in the linker) and SauI (site in the SPR-1 cDNA at position 292) and cloned into the BamHI/SauI sites of the plasmid pKS-A8O.

Expression in *E.coli*

SPR-1, SPR-2 and Sp1 cDNA expression vectors (see above) were transformed into *E.coli* strain BL21(LysS) (20). Cells were grown in LB medium to an OD_{600} of 0.5 and expression was induced by adding isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM. After 2 h at 37°C the bacteria were harvested and extracts prepared according to the procedure described for Sp1 (23).

Gel retardation experiments

Gel retardations assays were essentially performed as described (24,25). For the competition experiments, bacterial extracts containing SPR-1, SPR-2 and Sp1 were mixed and gel retardation assays performed with oligonucleotides containing the GC box and GT box binding sites, respectively. The GC box binding site was obtained by hybridization of the oligonucleotides 5′-AGC-TTCCGTTGGGGCGGGGCTTCACG -3′ and 5′-TCGACGTG-AAGCCCCGCCCCAACGGA-3′. The GT box binding site was obtained by hybridization of the oligonucleotides 5′-AGCTT-CCGTTGGGGTGTGGCTTCACG-3′ and 5′-TCGACGTGAA-GCCACACCCCAACGGA-3′.

Northern blot analysis

Total RNA from human cell lines and mouse organs was extracted by the guanidinium isothiocyanate/CsCl procedure (26), separated on 1.2% agarose gels containing 2.2 M deionized formaldehyde and blotted to nylon membranes (Pall Biodyne B). Prehybridization and hybridization was carried out as described (27) using ^{32}P labeled cDNA fragments as probes. The specific activities of the probes were approximately $5 \cdot 10^8$ cpm/μg. Hybridization probes for SPR-1, SPR-2 and Sp1 mRNA were as follows. SPR-1: 0.8 kb SauI-ScaI fragment (nucleotides 291 to 1115 in Fig. 2 B) or a 0.5 kb EcoRI fragment from pKS-A8O (containing nucleotides 2528 to 2971 shown in Fig. 2 B). SPR-2: 0.4 kb KpnI fragment from A3O (containing nucleotides 1270 to 1583 shown in Fig. 2 D) or a 180 bp EcoRV-NotI fragment from pKS-A3.1 (containing nucleotides 2169 to 2349). Sp1: 1.8 kb SstI fragment from plasmid pBS-Sp1-fl (provided by Dr. Tjian). Control hybridizations of 18S ribosomal RNA have been performed as described (28) using a ^{32}P labeled single stranded oligonucleotide probe that has the sequence 5′-ACGGTATCTG-ATCGTCTTCGAACC-3′.

RESULTS

Cloning of proteins which bind to the GT1 motif

To identify the protein(s) of Ishikawa cells which bind(s) to the GT1 motif of the uteroglobin promoter (Fig. 1) we decided to clone by recognition site screening their corresponding cDNA(s). Ishikawa cell cDNA expression libraries were generated by either random or oligo(dT) priming. The random primed library was screened for DNA binding proteins that recognize the GT1 motif. Initially, two cDNA clones of 2.8 kb (A8O) and 1 kb (A3O) in length have been obtained (Fig. 2 A and C). Sequence analysis revealed that these cDNAs encode two different proteins, designated henceforth SPR-1 and SPR-2.

To obtain a more complete cDNA for SPR-2, both Ishikawa cDNA libraries were screened with subfragments of the A3O clone. Sequence comparison of the clones obtained by this screening procedure with the A3O clone shows that A3O differs at the 3′ end from the corresponding nucleotide sequences of four other clones (A3.1, A3.11, A3.21 and A3.22). Northern blot analysis of RNA from Ishikawa cells with the 3′ end of the A3O clone and the corresponding sequence of the A3.1 clone reveals that the latter but not the A3O-specific probe specifically hybridizes with RNA (data not shown). Therefore, we assume that the 3′ end of A3O represents a cloning artifact generated during cDNA synthesis.

To obtain the complete open reading frame of SPR-1, ligation mediated PCR (21,22) was employed using a SPR-1-specific primer that hybridizes 180 nucleotides downstream of the 5′ end

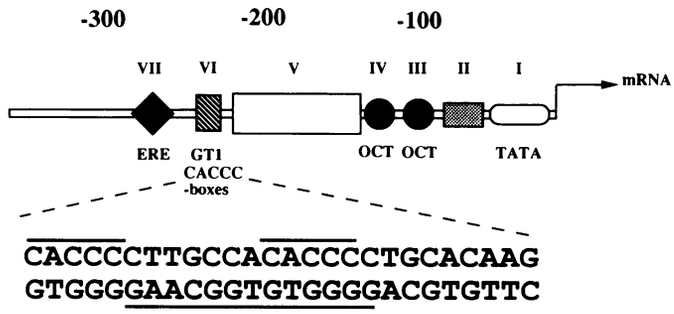


Figure 1. Structure of the rabbit uteroglobin promoter and nucleotide sequence of the GT1 motif. The uteroglobin promoter contains 7 mutation-sensitive regions (I–VII) (1,2). Abbreviations: TATA, TATA-box region; OCT, octamer factor binding site; ERE, estrogen responsive element.

of the A8O clone. By this approach, we have generated two clones (A8.68 and A8.90) which have A8O 5′ extensions (Fig. 2 A).

The SPR-1 cDNA sequence shown in Fig. 2 B is 2986 bp in length and has a potential AUG start codon followed by an open reading frame (ORF) of 784 amino acids. The calculated molecular weight of the product of the ORF is 81.9 kDa. The SPR-2 cDNA sequence shown in Fig. 2 D probably does not contain the complete open reading frame. The sequence is 3548 bp in length and the open reading frame codes for 697 C-terminal amino acids.

Structural features of SPR-1 and SPR-2

Inspection of the amino acid sequence, as deduced from the cDNA sequence, reveals that SPR-1 and SPR-2 have three zinc finger motifs of the Cys_2/His_2 type near the C-terminus. These structural motifs were first identified as nine tandemly repeated sequences in the RNA polymerase III transcription factor TFIIIA (29,30) and have been found later in many transcription factors from various species. The three zinc fingers of SPR-1 and SPR-2 are 90% identical to each other and to the zinc fingers of the transcription factor Sp1 originally isolated from HeLa cells (31) (Fig. 3). In addition, the amino acid sequences N-terminal of the zinc finger domain of SPR-1 and SPR-2 are also homologous to each other as well as to Sp1 (approximately 35% identity) and are composed of long serine/threonine- and glutamine-rich regions (Fig. 3 A) which have been identified for Sp1 as transactivation domains (23,32,33). Moreover, the amino acid sequences of SPR-1 and SPR-2 show 8 and 6 potential glycosylation sites (N-X-S/T), respectively, similar to the Sp1 amino acid sequence that has 10 potential glycosylation sites (31,33). Thus, it is likely that SPR-1 and SPR-2 are posttranslationally modified by glycosylation as it has been shown for Sp1 (34).

DNA binding specificity

The high degree of conservation of the zinc finger regions of Sp1, SPR-1 and SPR-2 (Fig. 3 B) prompts the question whether all three proteins recognize the same DNA motifs with similar affinities. The structure of the Sp1 DNA binding domain itself is not known but the crystal structure of the protein-DNA complex of another DNA binding protein containing three zinc fingers of the Cys_2/His_2 class, Zif268 (a mouse immediate early protein, also known as Krox-24) has been determined (35). The three zinc fingers of Zif268 are related to Sp1 and bind the 9−base-pair target sequence 5′-GCGTGGGCG-3′. These structural data and *in vitro* directed mutagenesis guided by similarities between the zinc fingers of transcription factors Sp1 and Krox-20 (36) (Krox-20 has the same DNA binding domain as Krox-24) have identified amino acids governing DNA-binding specificity. All amino acids that are putatively involved in DNA recognition are fully conserved between Sp1, SPR-1 and SPR-2. Therefore, it appears that Sp1, SPR-1 and SPR-2 may recognize identical DNA sequences. To test this, we have subcloned zinc finger containing regions of Sp1, SPR-1 and SPR-2 encoding 773, 263 and 778 C-terminal amino acids, respectively, in appropriate pET vectors (20), expressed them in *E.coli*, and employed gel retardation assays using oligonucleotides containing the central part of the GT motif of the uteroglobin promoter (5′-GGGGTGT-GG-3′) and a typical GC box (5′-GGGGCGGGG-3′) (37). Unique complexes are generated with SPR-1 and SPR-2 containing bacterial extracts using a ^{32}P labeled oligonucleotide containing the GT motif. Bacterially expressed full length Sp1 produces three different complexes. This phenomenon has already been observed previously (38). Probably, the two faster migrating complexes are due to proteolytic degradation products of Sp1 or perhaps result from internal translation start sites. Nevertheless, all complexes are competed identically with excess amounts of unlabeled oligonucleotides containing either the GT or the GC motif (Fig. 4 A) but not with oligonucleotides containing a random sequence or with poly (dI-dC) (data not shown). Therefore, it appears that Sp1, SPR-1 and SPR-2 bind the GT motif with identical affinities. In addition, the GC motif is recognized also identically by all three proteins but with an approximately 2-fold higher affinity than the GT box (Fig. 4 B).

Expression in human cell lines and mouse organs

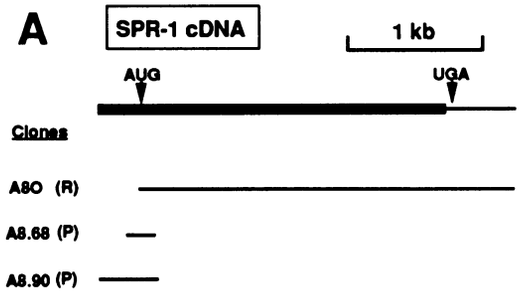
Northern blot analyses of total RNA from different human cell lines (Fig. 5 A) reveal that SPR-1 and SPR-2 mRNA are unique species of 7.5 kb and 5.0 kb. In comparison, Sp1 mRNA is 8.2 kb in length. Like Sp1 mRNA, SPR-2 mRNA is present in all cell lines tested and the relative amount of this mRNA differs only moderately. In contrast, SPR-1 mRNA concentration in these cell lines varies much more. It is highest in the two endometrial cell lines Ishikawa and MFE 296 and is not found in the lung epithelial cell line EPLC32M1.

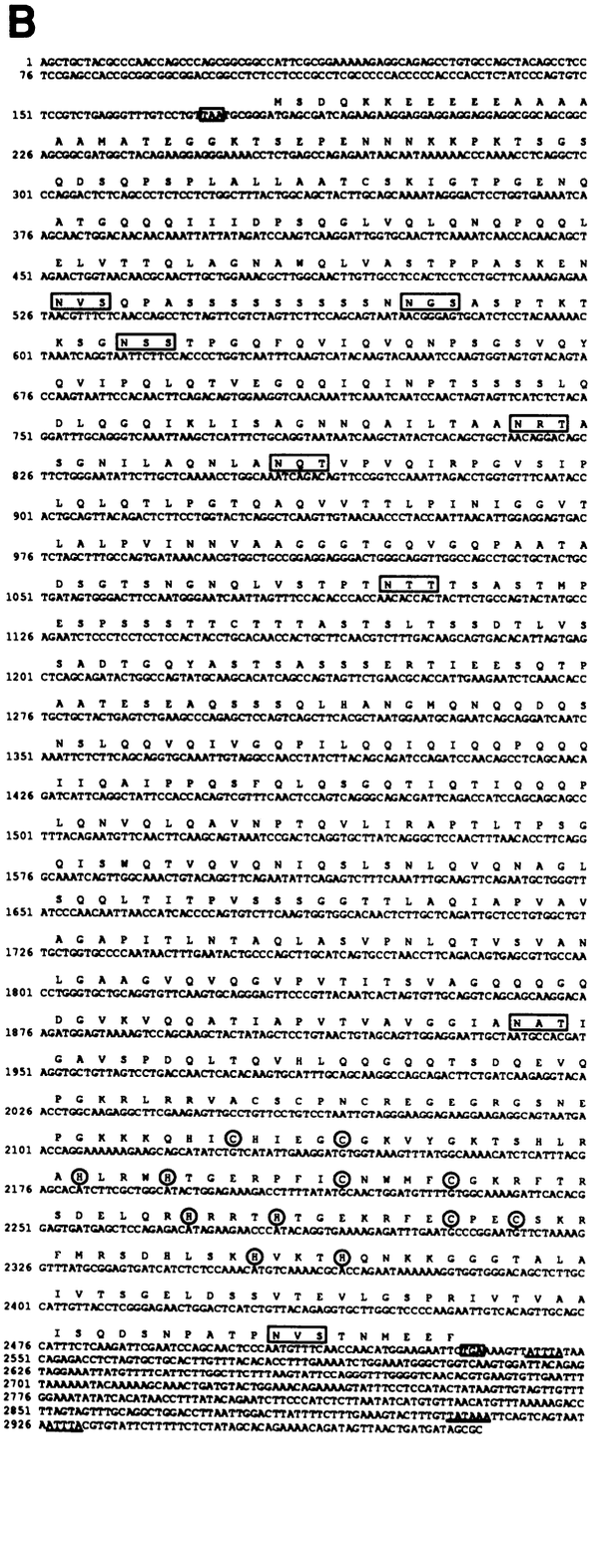
To address the question how SPR-1 and SPR-2 mRNA levels vary *in vivo*, we performed Northern blot analyses with total RNA from mouse tissues using human SPR-1, SPR-2 and Sp1 cDNA probes outside the zinc finger encoding regions. Except for RNA from testis, all three cDNAs specifically hybridize to unique mRNA species from mouse organs (Fig. 5 B). Sp1 and SPR-2 transcripts are present in all organs. The highest mRNA level is detectable in brain. Smaller SPR-2 and Sp1 transcripts appear exclusively in testis, suggesting that they arise from testis-specific processes. SPR-1 transcripts are abundantly present in brain but are barely detectable in other organs and tissues (salivary gland, liver, kidney, uterus, testis, seminal vesicle, lung, epididymis and esophagus). Quantitative evaluations (data not shown) show that the level of SPR-1 mRNA in brain is at least ten fold higher than in any other organ suggesting a brain-specific role of SPR-1.

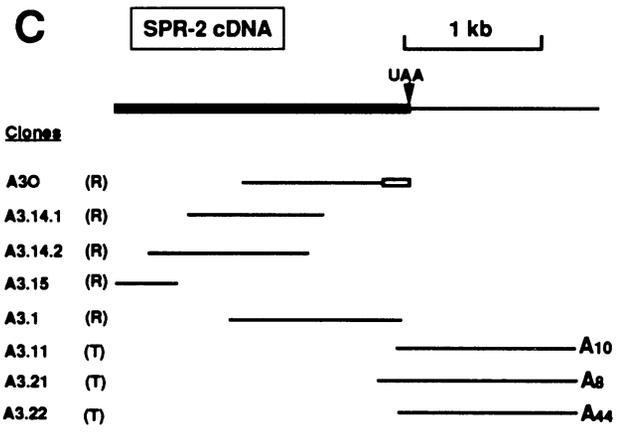
DISCUSSION

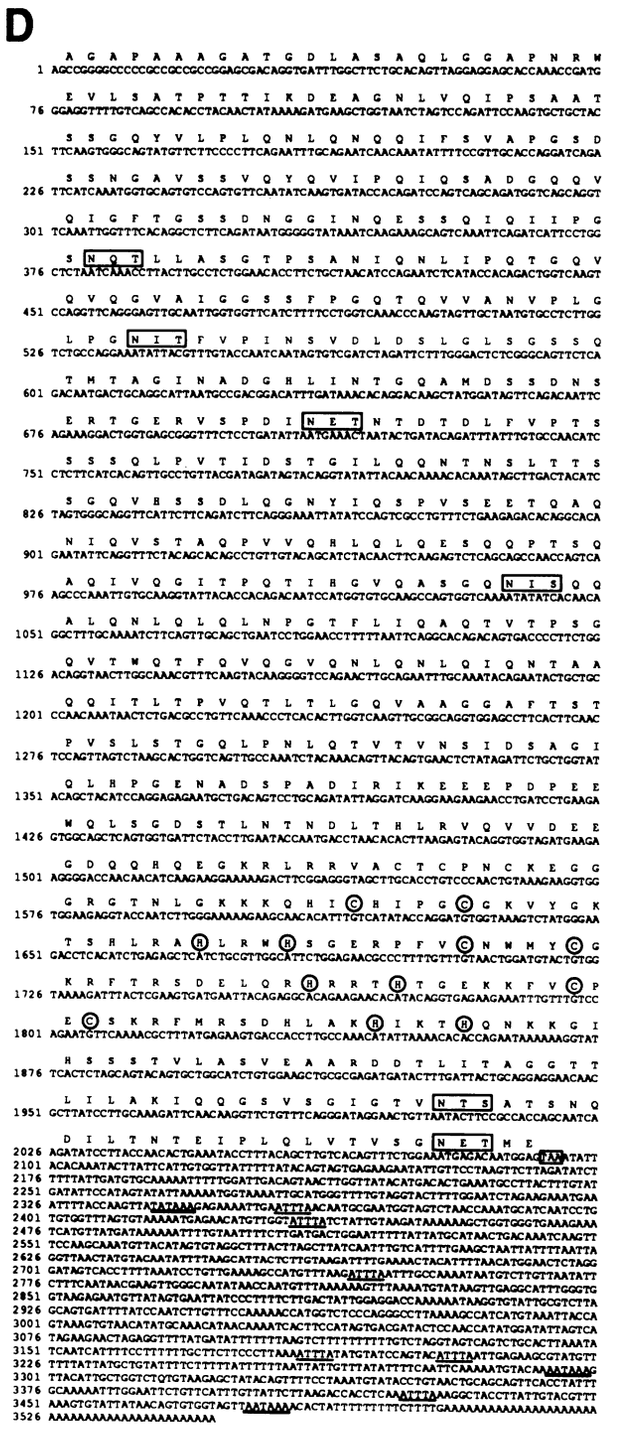
We have previously identified elements of the uteroglobin promoter that contribute to its activity in the endometrial cell line Ishikawa. One of these elements is located 230 bp upstream of the transcriptional start site and contains a GT-rich motif that is also present in the SV40 enhancer. This element in the context of the uteroglobin promoter enhances transcription of a reporter gene in Ishikawa but not in HeLa cells (1). Nevertheless, differences in the GT1 binding activity have not been detected between nuclear extracts from Ishikawa and HeLa cells (1 and unpublished results). One assumption to explain these results was that different proteins of similar size and binding properties might bind to this motif. The cloning by recognition site screening of SPR-1 and SPR-2 as members of a multigene family including Sp1 supports this idea. SPR-1 and Sp1 have calculated molecular weights of 81.9 kDa and 79.8 kDa, respectively. The open reading frame of the cDNA sequence of SPR-2 also indicates that this clone encodes a protein larger than 74 kDa.

Surprisingly, SPR-1 and SPR-2 transcripts are also present in HeLa cells, although their relative concentrations in this cell line differ moderately from Ishikawa cells. If the presence of the SPR-1 and SPR-2 mRNAs in both cell lines reflects the protein levels, additional mechanisms must be postulated to explain the


A
SPR-1 cDNA
1 kb
AUG
UGA
Clones
A8O (R)
A8.68 (P)
A8.90 (P)


B


C
SPR-2 cDNA
1 kb
UAA
Clones
A3O (R)
A3.14.1 (R)
A3.14.2 (R)
A3.15 (R)
A3.1 (R)
A3.11 (T)
A3.21 (T)
A3.22 (T)
A10
A8
A44


D

finding that the GT1 motif of the uteroglobin promoter is active in Ishikawa cells but not in HeLa cells. One possibility is that SPR-1 and SPR-2 act in combination with other, endometrium-specific factors (see below). Alternatively, even more members of this new multigene family exist that are specifically expressed in Ishikawa cells. This is not unlikely as the recognition site screening has not revealed clones for Sp1, although Sp1 transcripts are present in Ishikawa cells in amounts comparable to those for SPR-1 and SPR-2 mRNA.

SPR-2 and Sp1 transcripts are also abundant in uterus and lung and, therefore, may be involved in transcription of the uteroglobin gene *in vivo*. Unexpectedly, SPR-1 mRNA is abundant in brain but barely detectable in other organs. Nevertheless, this finding does not exclude that SPR-1 is also involved in uteroglobin gene transcription. Uteroglobin is only transcribed in some specialized epithelial cells of certain tissues. For instance, in the endometrium, uteroglobin is expressed in the glandular epithelial

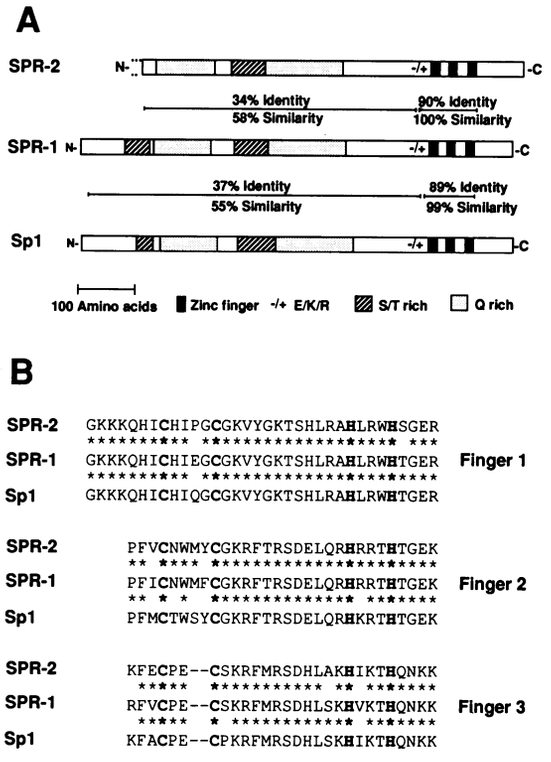


Figure 3. Primary structure features of SPR-1 and SPR-2 in comparison with Sp1. (**A**) Each differently filled box refers to the region containing a preponderance of one or several types of amino acids, which are indicated at the bottom using the single letter code. The values for identities and similarities were obtained by protein sequence alignments using the Bestfit program of the GCG Sequence Analysis Software Package (Genetics Computer Group, Madison). (**B**) Protein sequence alignment of the zinc finger domains of SPR-1, SPR-2 and Sp1. Asterisks represent amino acid identities. The cysteine and histidine residues which are probably involved in zinc binding are depicted with bold letters.

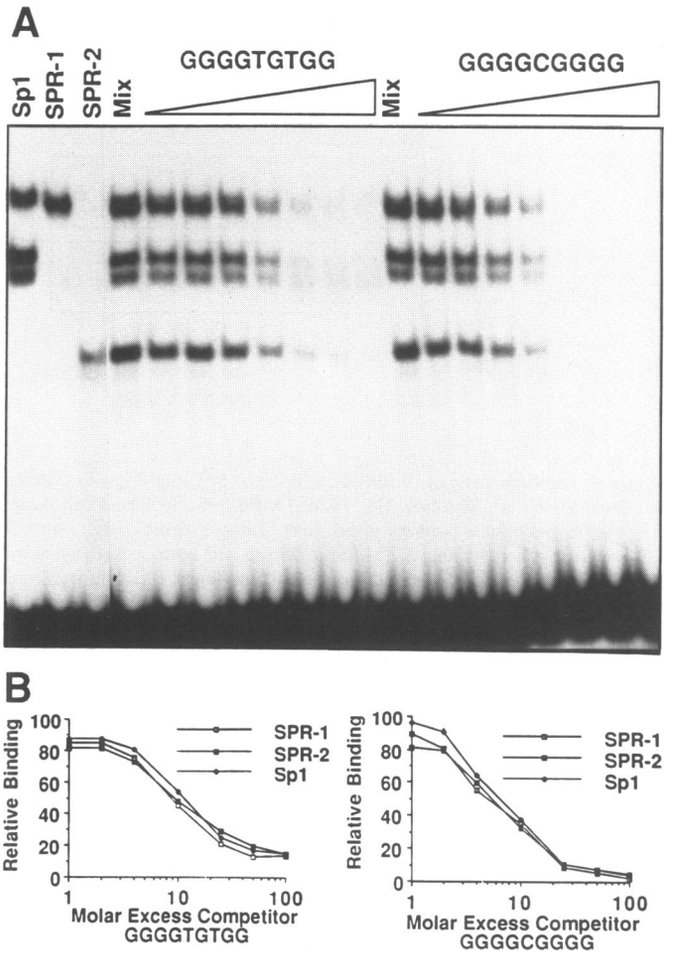


Figure 4. Binding specificity of SPR-1, SPR-2 and Sp1. (**A**) Gel retardation analysis of the interactions of SPR-1, SPR-2 and Sp1 with their target sequences. Bacterial extracts used containing SPR-1, SPR-2, Sp1 or a mixture of all three proteins (mix) are indicated at the top of the figure. All reactions contained 0.1 ng of labeled GT oligonucleotide and various amounts (1 to 100-fold molar excess) of non-labeled GT and GC oligonucleotide as indicated. The nucleotide sequence of the central part of the oligonucleotides is shown above the autoradiograph. (**B**) Quantitative evaluation of competition assays. The radioactivity in the complexes (arbitrary value of 100 in the complexes without specific competitor DNA) is plotted against the molar excess of competitor DNA. To distinguish between the slowest migrating complex of Sp1 and the complex of SPR-1 in the 'mix' reactions quantitative evaluations of competition experiments using single extracts only have been performed.

Figure 2. Structure and nucleotide sequences of the cDNA clones encoding human SPR-1 and SPR-2. (**A** and **C**) Schematic representation of the overlapping Ishikawa cell cDNA clones from the random primed λgt11 expression library (R), the oligo(dT) primed library (T), and the ligation mediated PCR products (P). The bold lines refer to the open reading frames. The first AUG codon of SPR-1 and the stop codons of SPR-1 and SPR-2 are indicated. The name of each clone is given on the left side of the diagram. The clone A3O differs at the 3′ end from the corresponding nucleotide sequences of clones A3.1, A3.11, A3.21 and A3.22, probably due to an artifact during cDNA synthesis. (**B** and **D**) Nucleotide and deduced amino acid sequences of SPR-1 (B) and SPR-2 (D). Numbers on the left correspond to the DNA sequence. The stop codons at the end of the open reading frame of both cDNAs and the in frame stop codon in the 5′ untranslated region of SPR-1 are boxed. The positions of potential polyadenylation and mRNA destabilization signals (43) in the 3′ untranslated regions are underlined bold and plain, respectively. The derived amino acid sequence is shown above the nucleotide sequence using the single letter code. The cysteine and histidine residues which are putatively involved in zinc binding are circled. Boxes indicate potential N-glycosylation sites (N-X-S/T).

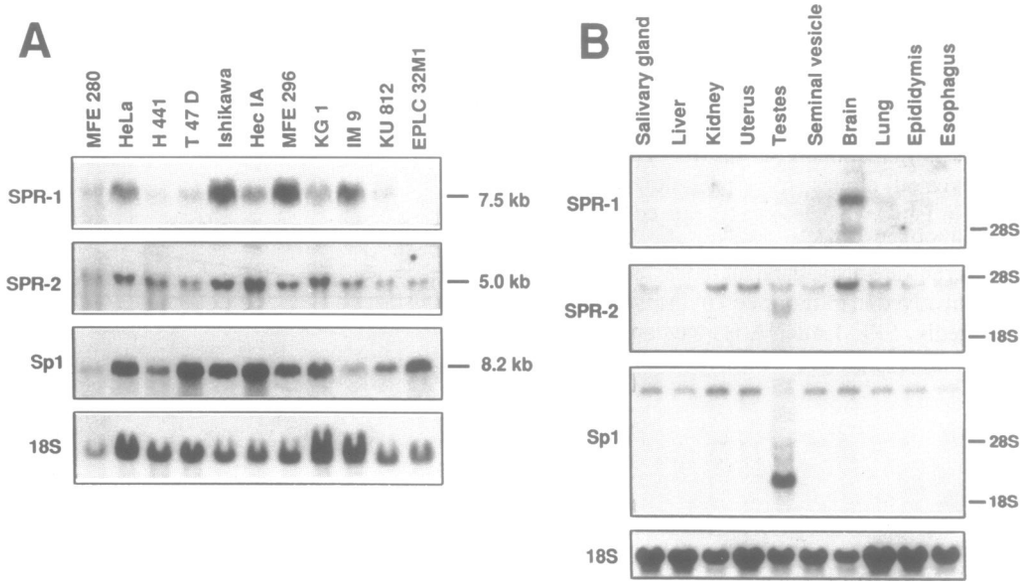


Figure 5. Northern analysis of SPR-1, SPR-2 and Sp1 transcripts in different human cell lines (**A**) and mouse tissues (**B**). Total RNA (15 μg) from endometrial cell lines (MFE 280, Ishikawa, Hec 1A and MFE 296), lymphoid cell lines (KG1, IM9 and KU812), lung cell lines (H441 and EPLC32M1), HeLa cells, T47D cells and mouse tissues (salivary gland, liver, kidney, uterus, testis, seminal vesicle, brain, lung, epididymis and esophagus) was subjected to electrophoresis on 1.2% agarose gels containing 2.2 M formaldehyde, and transferred to nylon hybridization membranes. The filters were successively hybridized with cDNA fragments of SPR-1, SPR-2 and Sp1 that did not contain zinc finger encoding regions. For the human RNA blot commercially available RNA size markers were used. The sizes of the human transcripts and the migration positions of the 28S and 18S ribosomal RNAs are indicated on the right. The control hybridization of the filters using an 18S ribosomal RNA specific probe is shown at the bottom.

cells and in the lung only a minor cell fraction, the Clara cells, express uteroglobin (39). For the Northern blot analysis of SPR-1 mRNA we have used total RNA from the whole uterus and lung and have obtained only weak signals. Immunohistochemical studies using sections of these organs have to be performed to analyze the expression pattern of SPR-1 within an organ to identify a potential cell type-specific expression pattern that possibly correlates with the expression of the uteroglobin gene.

Binding sites for SPR-1, SPR-2 and Sp1 are present in many enhancers and promoters. Thus, it is very likely that the members of this new gene family are involved in the expression of many genes. All three proteins not only share their DNA binding domain with each other but also glutamine-rich regions which have been identified for Sp1 as transactivation domains (23). Therefore, SPR-1 and SPR-2 probably also act as transactivators. But how might SPR-1, SPR-2 and Sp1 exert specific functions? It has been reported that stimulation of transcription by Sp1 may be dependent on the promoter context, as Sp1 sites placed upstream of some promoters do not allow Sp1 to function synergistically (40). This raises the intriguing possibility that SPR-1, SPR-2 and Sp1 function promoter-specifically by interacting with different transcription factors. This issue cannot be easily clarified because our Northern blot analyses indicate that most cell lines contain all three proteins. The *Drosophila* SL2 cell line transcription system which has been used to characterize Sp1 may not be appropriate to address this question either, because other transcription factors that could be involved in the activity of a certain promoter might not be present in this cell line. *In vitro* transcription assays using nuclear extracts depleted of SPR-1, SPR-2 and Sp1 may have to be established to address these questions.

Sp1 can stimulate transcription synergistically from two or multiple binding sites (adjacent or far distant) in the *Drosophila* cell line SL2. For this synergistic action, the domain C-terminal to the zinc finger region (domain D of Sp1) is required (40). In contrast to other domains, the amino acid sequence of this domain is not conserved between Sp1, SPR-1 and SPR-2. Thus, one could speculate that these parts of the proteins may play a crucial role for the combinatorial action with other transcription factors.

In conclusion, our results imply that the presence of GC or GT boxes in many promoters and enhancers does not necessarily mean that Sp1 is the sole transcription factor involved in expression of these genes. At least two other homologous factors, SPR-1 and SPR-2 that have highly conserved Sp1-like zinc finger regions also bind to GT as well as GC boxes with identical affinities and very likely act as transcriptional regulators as well. It appears that like other control elements, such as AP1 and CRE sites which bind families of transcription factors of at least six and seven members, respectively, (41,42), GT and GC boxes are also recognized by a family of factors. Therefore, the discovery of SPR-1 and SPR-2 introduces a new complexity in the combinatorial possibilities of transcriptional regulation exerted through GT and GC boxes. Functional investigations will be necessary to understand the interplay between Sp1, SPR-1 and SPR-2. Further examination of the cellular and developmental specificity of the expression pattern of SPR-1, SPR-2 in comparison to Sp1 as well as gene disruption experiments with the corresponding mouse genes will also shed more light on the physiological role of these new proteins.

After preparation of the manuscript the cloning of Sp1 related genes have been reported [Imataka, H. *et al.* (1992) *EMBO J.*,

11, 3663–3671; Kingsley,C. and Winoto,A. (1992) *Mol. Cell. Biol.*, **12**, 4251–4261]. It appears that the Sp3 clone published by Kingsley and Winoto is identical to SPR-2. To simplify the nomenclature we propose to rename SPR-2 and SPR-1. SPR-2 should be called Sp3 and SPR-1 should be called Sp4.

ACKNOWLEDGEMENTS

We thank W.Lorenz for excellent technical assistance and V.Rönicke for help during screening. Drs.E.Pascal and R.Tjian kindly provided us with Sp1 cDNA clones. Drs.M.Kalff-Suske, J.Klug and M.Truß are gratefully acknowledged for critically reading the manuscript. This work was supported by a grant of the Deutsche Forschungsgemeinschaft.

REFERENCES

1. Misseyanni,A., Klug,J., Suske,G. and Beato,M. (1991) *Nucleic Acids Res.*, **19**, 2849–2859.
2. Suske,G., Lorenz,W., Klug,J., Gazdar,A.F. and Beato,M. (1992) *Gene Expression*, **2**, in press.
3. Jones,N.C., Rigby,P.W.J. and Ziff, E.B. (1988) *Genes Dev.*, 2, 267–281.
4. Zenke,M., Grundström, T., Matthes, H., Wintzereith, M., Schatz, C., Wildeman, A. and Chambon, P. (1986) *EMBO J.*, **5**, 387–397.
5. Xiao,J.-H., Davidson,I., Macchi,M., Rosales,R., Vigneron,R., Staub,A. and Chambon,P. (1987) *Genes Dev.*, **1**, 794–807.
6. Becker, P. B., Gloss, B., Schmid, W., Strähle, U. und Schütz, G. (1986), *Nature*, **324**, 686–688.
7. Strähle,U., Schmid,W. and Schütz,G. (1988) *EMBO J.*, 7, 3389–3395.
8. Danesch,U., Gloss, B., Schmid, W., Schütz,G. and Renkawitz,R. (1987) *EMBO J.*, **6**, 625–630.
9. Schüle,R., Muller, M., Otsuka-Murakami,H. and Renkawitz,R. (1988) *Nature*, **332**, 87–90.
10. Myers,R.M., Tilly,K. and Maniatis,T. (1986) *Science*, **232**, 613–618.
11. Yu,C.-Y., Motamed,K., Chen, J., Bailey,A.D. and Shen, C.-K.L. (1991) *J. Biol. Chem.*, **266**, 8907–8915.
12. Gubler, U. and Hoffmann, B.J. (1983) *Gene*, **25**, 263–269.
13. van Doorn,L.-J., van Belkum,A. and Kos,T. (1989) *Nucleic Acids Res.*, **17**, 9496.
14. Nishida,M., Kasahara,K., Kaneko,M. and Iwasaki,H. (1985) *Acta Obstet. Gynecol. Jpn.*, **37**, 1103–1111.
15. Aguan,K., Kusano,T., Suzuki, N. and Kitagawa,Y. (1990) *Nucleic Acids Res.*, **18**, 1071.
16. Singh,H., LeBowitz,J.H., Baldwin,A.S. and Sharp,P.A.(1988) *Cell*, **52**, 415–423.
17. Vinson,C.R., LaMarco,K.L., Johnson,P.F., Landschulz,W.H. and McKnight,S.L. (1988) *Genes Dev.*, **2**, 801–806.
18. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 5463–5467.
19. Hattori,S. and Sakaki,Y. (1986) *Anal. Biochem.*, **152**, 232–238.
20. Studier,F.W., Rosenberg,A.H., Dunn,J.J. and Dubendorff,J.W. (1989) *Methods Enzymol.*, **185**, 60–89.
21. Mueller,P.R. and Wold,B. (1989) *Science*, **246**, 780–786.
22. Pfeifer, G.P., Steigerwald,S., Mueller, P.R. and Riggs, A.D. (1989) *Science*, **246**, 810–813.
23. Courey,A.J. and Tjian,R. (1988) *Cell*, **55**, 887–898.
24. Fried,A. and Crothers,D.M. (1981) *Nucleic Acids Res.*, **9**, 6505–6525.
25. Garner,M.M. and A. Revzin (1986) *TIBS*, **11**, 395–396.
26. Chirgwin,J.M., Przybyla,A.E., MacDonald,R.J. and Rutter,W.J. (1979) *Biochemistry*, **18**, 5294–5299.
27. Hagen,G., Wolf,M., Katyal,S.L., Singh,G., Beato,M. and Suske,G. (1990) *Nucleic Acids Res.*, **18**, 2939–2946.
28. Szyf,M., Milstone,D.S., Schimmer,B.P., Parker, K.L. and Seidman,J.G. (1990) *Mol. Endocrinol.*, **4**, 1144–1152.
29. Brown,R.S., Sander,C. and Argos,P. (1985) *FEBS Lett.*, **186**, 271–274.
30. Miller,J., McLachlan,A.D. and Klug,A. (1985) *EMBO J.*, **4**, 1609–1614.
31. Kadonaga,J.T., Carner,K.R., Masiarz,F.R. and Tjian,R. (1987) *Cell*, **51**, 1079–1090.
32. Courey,A.J., Holtzman,D.A., Jackson,S.P. and Tjian,R. (1989) *Cell*, **59**, 827–836
33. Kadonaga,J.T., Courey,A.J., Ladika,J. and Tjian,R. (1988) *Science*, **242**, 1566–1570.
34. Jackson,S.P. and Tjian,R. (1988) *Cell*, **55**, 125–133.
35. Pavletich,N.P.and Pabo,C.O. (1991) *Science*, **252**, 809–817.
36. Nardelli,J., Gibson,T.J., Vesque,C. and Charnay,P. (1991) *Nature*, **349**, 175–178.
37. Kadonaga,J.T., Jones,K.A. and Tjian,R. (1986) *TIBS*, **11**, 20–23.
38. Chavrier,P., Vesque,C., Galliot,B., Vigneron,M., Dollé,P., Duboule,D. and Charnay,P. (1990) *EMBO J.*, **9**, 1209–1218.
39. Sandmöller,A., Voss,A.K., Hahn,J., Redemann-Fibi,B., Suske,G. and Beato,M. (1991) *Mech. Dev.*, 34, 57–68.
40. Pascal,E. and Tjian,R. (1991) *Genes Dev.*, **5**, 1646–1656.
41. Kouzarides,T. and Ziff,E. (1989) *Cancer Cells*, **1**, 71–76.
42. Hai,T., Liu,F., Coukos,W.J. and Green,M.R. (1989) *Genes Dev.*, **3**, 2083–2090.
43. Malter,J.S. (1989) *Science*, **246**, 664–666.