Transcription of genes encoding pregnancy-specific glycoproteins is regulated by negative promoter-selective elements

Graciela M. PANZETTA-DUTARI, Nicolás P. KORITSCHONER, José L. BOCCO, Rodrigo NORES, Catherine I. DUMUR and Luis C. PATRITO¹

Departamento de Bioquímica Clínica, Facultad de Ciencias Químicas, Universidad Nacional de Córdoba, Ala Oeste, Ciudad Universitaria, 5000 Córdoba, Argentina

The human pregnancy-specific glycoprotein (*PSG*) genes comprise a family of 11 highly conserved members whose expression is maximal in placental cells and marginal in other cell types. We have investigated here the molecular basis of *PSG* regulation by analysing a large regulatory region of the *PSG-5* gene in cells that do and do not express these genes. The promoter region (–254 to –43), which does not contain a TATA-box, large GC-rich sequences or a classical initiator, was active in all cell types analysed. Additional upstream sequences up to position –3204 repressed promoter activity. Two independent repressor regions were identified and found to operate effectively in HeLa, COS-7 and HTR8/SVneo placental cells. More significantly, these negatively acting modules failed to repress a heterologous TATA-

containing thymidine kinase promoter. Detailed transcriptional and DNA–protein analyses of the proximal repressor region (–605 to –254) revealed the presence of both negative and positive *cis*-acting elements. Disruption of the repressive functions resulted in an enhanced transcription of the reporter constructs. In conclusion, these results demonstrate that *PSG-5* gene transcription is highly repressed by promoter-selective negative regulatory regions and the relief of repression allows enhanced *PSG-5* gene transcription irrespective of the cell type. Furthermore, our findings suggest that *PSG* genes are expressed mainly through a derepression mechanism.

Key words: gene control, repression, TATA-boxless.

INTRODUCTION

The placenta is an essential organ for mammalian reproduction that links maternal and fetal compartments. Placental cytotrophoblastic and syncytiotrophoblastic cells are highly proliferative and invasive. In addition, these cells display a versatile metabolic activity, represented mainly by the synthesis of hormones, cytokines, growth factors and their receptors [1], as well as by a number of pregnancy-specific proteins such as the pregnancy-specific glycoproteins (PSGs). These proteins are the major placental proteins found in maternal circulation during pregnancy, reaching concentrations of approx. 200 µg/ml at term [2]. Although the function of PSGs remains elusive, they seem to be essential for the maintenance of normal pregnancy because low levels are associated with poor outcome of pregnancy [3,4]. PSGs comprise a family of highly similar polypeptides encoded by 11 transcriptionally active genes [5], clustered on chromosome 19q13.2 along with the carcinoembryonic antigen (*CEA*) gene family [6]. The PSGs share significant nucleotide and amino acid sequence similarity with the *CEA* members. In spite of their high similarity, which includes approx. 200 nt upstream from the translational start site, they show a different pattern of gene expression. *CEA* members are expressed mainly in epithelial cells, granulocytes and lymphocytes [7]. In contrast, high levels of PSG proteins are limited to the syncytiotrophoblast cells and low levels are found in fetal liver, myeloid cells, trophoblastic tumours and in some cell lines such as HeLa and placental HTR8/SVneo cells [8–12] (see below).

At present, our understanding of the factors that regulate the

expression of *PSG* genes is in its early stages. The biosynthesis of PSGs is regulated mainly at the transcriptional level, increasing their expression during placental development [8]. The promoters of the *PSG* genes show an interesting structure because they lack a TATA-box, a pyrimidine-rich initiator element and large GC regions [13,14]. We have previously demonstrated that the *PSG-5* promoter contains a GC-box that acts as a core promoter element (CPE) (CCCCACCC) and is localized 30 bp upstream of the major transcription start site, resembling an initiator element. This sequence contributes most of the promoter activity in a variety of cell lines [15]. Furthermore, this crucial CPE element is recognized *in vitro* by a Krüppel-like transcription factor, CPBP, that is highly expressed in human placenta at term [16] and throughout mammalian gestation [17]. Similar CPE sequences described in the promoter of other *PSG/CEA* members have been associated directly with the interaction of transcription factors related to Sp1, USF and AP2 [18–21].

The unique promoter features, restricted expression pattern and relatively recent evolution [22] of *PSG* genes have prompted considerable interest in identifying some elements involved in their control of expression. To obtain insight into the mechanisms that turn on *PSG* expression we have characterized an extensive 5′-flanking *PSG-5* sequence. Our findings indicate that *PSG-5* gene transcription is highly repressed in non-placental cells and also in placenta-derived HTR8/SVneo cells. Disruption of the repressive activity permits an enhanced expression of the gene, even in COS-7 cells, in which the endogenous *PSG* genes are not transcribed. In addition, our results suggest the existence of a specific derepression mechanism underlying the transactivation of *PSG* genes.

Abbreviations used: CAT, chloramphenicol acetyltransferase; *CEA*, carcinoembryonic antigen; CPE, core promoter element; DRR, distal repressor region; EMSA, electrophoretic mobility-shift assay; PRR, proximal repressor region; *PSG*, pregnancy-specific glycoprotein; tk, thymidine kinase.

¹ To whom correspondence should be addressed (e-mail lpatrito@fcq.unc.edu.ar).

The nucleotide sequence data reported will appear in DDBJ, EMBL and GenBank® Nucleotide Sequence Databases under the accession number AF202886.

Table 1 Oligonucleotides used in this study

FP1:	5'TTGAAAGAGGGCAAAGGGCGTGAATG 3' CTTTCTCCCGTTTCCCGCACTTAC
mutA:	5'TTGAAAatctagAAAGGGCGTGAATGTG 3' CTTTtagatcTTTCCCGCACTTACAC
mutB:	5'TTGAAAGAGGGCAAAGGGatcagcTGTG 3' CTTTCTCCCGTTTCCctagtcgACAC
FP3:	5'GCATCACTGTGACAGGGACCGTTGCAGCTT 3' GTGACACTGTCCCTGGGCAACGTCGAA
FP3mut:	5'GCATCACTGTGACAacgcgctCGTTGCAGCTT 3' GTGACACTGTtgcgcgcaGCAACGTCGAA
NS1:	5'AAGAGGCAATAATAAGGAAAT 3' and its complement

EXPERIMENTAL

Plasmid constructs

UBCAT contains a −254 to −43 fragment relative to the translation start point of the *PSG-5* gene inserted upstream of the chloramphenicol acetyltransferase (CAT) gene in the pBLCAT3 vector [23]. PBCAT contains a −605 to −43 fragment and SBCAT contains a −3204 to −43 fragment inserted in pBLCAT3. Construction of these plasmids has been described previously [13]. Construct SBCAT served as the parental vector in the subsequent construction of SBΔStuCAT and SBΔBstECAT. They were prepared by excision of the internal fragments −706 to −254 and −2360 to −878 by digestion with *StuI* and *BstEII* respectively, followed by religation with T4 DNA ligase. Construct −421CAT was obtained from SBCAT by the excision of the 5'-flanking sequence between −3204 and −421 by digestion with *SalI* and *BstXI*, followed by filling with T4 DNA polymerase and blunt-end ligation with T4 DNA ligase. Constructs −457CAT, −524CAT, −524CATmFP3 and −567CAT were obtained by cloning the fragments amplified with the respective 5' primers and a common 3' primer into the pGEM-T vector (Promega), followed by excision with appropriate enzymes and ligation into pBLCAT3. Plasmid iSBCAT contains the sequence of the first *PSG-5* intron inserted upstream of the 5' end of the −3204 to −43 fragment in construct SBCAT. The intron sequence (+65/+1299) was amplified with two specific primers, by using the genomic SG1 clone as template [13]. The amplified product was ligated into the pGEM-T vector (Promega), excised with appropriate restriction enzymes and ligated to the SBCAT plasmid. Clone iUBCAT was obtained from iSBCAT by excision of the −3204 to −254 fragment by digestion with *SalI* and *StuI*, filling in and blunt-end religation with T4 DNA ligase.

A p2 plasmid series was constructed by the insertion of different fragments of the *PSG-5* regulatory region upstream of the thymidine kinase (tk) promoter in the pBLCAT2 vector [23]. The *PstI*–*BamHI* fragment (−605 to −43) cloned in pUC18 was excised by digestion with *BamHI* and *HindIII* (present in the polylinker of pUC18) and inserted into the *BamHI*–*HindIII* sites of the pBLCAT2 vector, thereby creating construct p2PB. Plasmid p2PU was prepared from p2PB by deletion of the −257 to −43 region by digestion with *StuI* and *BamHI*, filling in with Klenow enzyme and ligating with T4 DNA ligase. Plasmid p2BstE was obtained by ligating the blunt-ended *BstEII* (−2360 to −878) fragment into pBLCAT2 digested with *SalI* and blunt-ended with Klenow enzyme.

Plasmids PBmutA and PBmutB, and also PBmFP3, were generated by a four-primer PCR-mediated mutagenesis of the

FP1 and FP3 sequences respectively. Constructs differed from the parental PBCAT plasmid only in the FP1 or FP3 mutated sequences. All constructs were checked by sequencing with the chain-termination method [24].

Cell culture, transfections and CAT assays

COS-7 and HeLa cells were grown in Dulbecco's modified Eagle's medium supplemented with streptomycin (0.1 mg/ml), penicillin (100 i.u./ml) and 5% or 10% (v/v) fetal calf serum respectively. HTR8/SVneo placental cells were grown in RPMI-1640 medium supplemented with streptomycin (0.1 mg/ml), penicillin (100 i.u./ml) and 10% (v/v) fetal calf serum. COS-7 and HeLa cells were transfected by calcium phosphate co-precipitation [25] with 5 or 10 µg of recombinant DNA species, 1 µg of β-galactosidase expression plasmid (pCH110; Pharmacia) and carrier DNA up to 10 or 15 µg respectively. HTR8/SVneo cells were transfected with the LipofectAMINE reagent (Gibco BRL). After 24 or 48 h, protein extracts were prepared and assayed for CAT activity as described previously [26]. The amounts of cell extracts employed were normalized to equivalent β-galactosidase activity units. Transfections were conducted a minimum of three times with different purifications of each construct. The percentage acetylation of chloramphenicol was determined by TLC followed by scintillation counting.

DNase I footprinting and electrophoretic mobility-shift assays (EMSAs)

Nuclear extracts for both footprinting and EMSA were prepared from human term placenta by the method of Dignam et al. [27], and from COS-7 and HeLa cells by the method of Finocchiaro et al. [28]. Protein concentrations were determined by the method of Bradford [29]. For DNase I footprinting analysis, the *PstI*–*NcoI* (−605 to −415) and *NcoI*–*StuI* (−415 to −254) *PSG-5* gene fragments, cloned in pUC18, were excised with appropriate restriction enzymes, purified by PAGE and radiolabelled with fill-in reactions with [α -³²P]dATP and Klenow enzyme [30]. The DNase I protection assays were performed essentially as described [31]. EMSA experiments were performed with ³²P-labelled oligonucleotides. The probes were incubated with 5–10 µg of protein extracts in a total volume of 20 µl that contained 10 mM Hepes, pH 7.9, 80 mM KCl, 0.1 mM EDTA, 10% (v/v) glycerol, 0.5 mM dithiothreitol, 0.5 mM PMSF, 1 µg of poly[d(I-C)] and 0.5 µg of denatured salmon sperm DNA. For competition experiments, the protein fractions were preincubated (10 min on ice) with the appropriate unlabelled oligonucleotide before the addition of the labelled probe. After incubation on ice for 15 min, binding reactions were subjected to 5% (w/v) PAGE in 0.25 × TBE buffer and detected by autoradiography.

Oligonucleotides primers, probes and competitors

The sequences of the oligonucleotides used were as shown in Table 1.

RESULTS

PSG-5 transcription is repressed by two negative elements, irrespective of the cell type

We have previously reported the cloning of the human *PSG-5* gene and the sequence of 1190 nt upstream of the translational start site [13]. Here we extend the 5' regulatory sequence up to position −3204. DNA sequencing and database searches revealed an almost perfect match to a recently submitted sequence

Figure 1 Identification of transcriptional regulatory regions of the human *PSG-5* promoter

(A) Relative CAT activities of different *PSG-5* promoter constructs and pBLCAT2 plasmid in HeLa, COS-7 and HTR8/SVneo cells. A schematic presentation of the reporter constructs is shown on the left. The numbers represent the nucleotide positions relative to the translation initiation site. A cross-hatched box denotes the sequence of the first intron. Also indicated are relevant restriction endonuclease cleavage sites. Relative CAT activities were determined and are shown as percentages of the activity of the UBCAT construct, which was arbitrarily set to 100%. (B) Transcriptional activity of the PRR and part of the DRR of *PSG-5* on the heterologous tk promoter. A schematic presentation of the CAT reporter constructs containing the tk promoter fused to the PRR and part of the DRR is shown on the left. Relative CAT activities of the constructs are presented as percentages of the CAT activity of the tk promoter alone assessed in HeLa and COS-7 cells. Results are means $\pm$ S.D. for at least three independent experiments. CAT assays were normalized to the β -galactosidase activity used as an internal control for transfection efficiency.

corresponding to a 2.5 Mb region in 19q13.2 containing a clustered *CEA/PSG* gene family. The sequence reported in the present paper is included between positions 37900 and 35891 of the aforementioned sequence (accession number AC005392).

To obtain insight into the mechanisms that trigger the expression of *PSG* genes, the transcriptional activity of an extensive 5'-flanking region of the *PSG-5* gene was analysed. To this end, transient transfection assays were performed in HeLa cells, which produce low levels of *PSGs*, and in COS-7 cells, which do not express *PSG* genes (Figure 1A). In both cell types, maximum promoter activity was exhibited by the UBCAT construct. The inclusion of further upstream sequences up to -605 and -3204 resulted in CAT activities that were approx. 40% and 15% respectively of that of UBCAT. The presence of these repressor regions was further demonstrated with constructs SBΔStuCAT and SBΔBstECAT, both of which exhibited low CAT activities. These results delineate a proximal repressor region (PRR, position -605 to -254) and a distal repressor region (DRR, position -3204 to -706) that down-regulate the *PSG-5* pro-

motor. In addition, activating sequences capable of overriding the silencing activity were also absent from the 1234 bp sequence of the first *PSG-5* intron (Figure 1A, constructs iSBCAT and iUBCAT).

To examine further the potential relevance of the repressor regions to the physiological regulation of *PSG-5*, the functional activity of *PSG-5* constructs was analysed in HTR8/SVneo placental cells [33], in which *PSG* expression has been confirmed (R. Nores, G. Panzetta-Dutari and L. C. Patrino, unpublished work). In this cell background, the CAT activities of the PBCAT and SBCAT constructs decreased to approx. 45% and 10% respectively of that of UBCAT (Figure 1A). Additionally, the activity of the reference tk promoter compared with that of the UB (-254 to -43) promoter was similar in HeLa, COS-7 and HTR8/SVneo cells (Figure 1A). Thus the UB promoter fragment contains essential DNA sequences that are necessary to support the transcription of the *PSG-5* gene and it is down-regulated by at least two negative DNA regulatory elements in the three cell types analysed.

Figure 2 Functional dissection of the PRR of the *PSG-5* gene

Various *PSG-5*-CAT constructs were created by deletion of the PBCAT construct. The numbers indicate the relative nucleotide positions from the translation start codon at the 5' end of each individual construct. All these constructs had a common 3' end at -43. Constructs were introduced into HeLa and COS-7 by transient transfection. Relative CAT activities, normalized to the β -galactosidase activity used as an internal control, were determined and are shown as percentages of the repression or activation of the UBCAT construct. Results are means $\pm$ S.D. for at least three independent experiments.

PRR is promoter-selective

Most of the silencers or negative regulatory regions identified so far down-regulate the expression driven by heterologous promoters such as the tk promoter [34]. Therefore, to characterize the silencing function of the PRR, its sequence was inserted upstream of the tk promoter in the pBLCAT2 vector. The PRR did not inhibit tk promoter activity but increased CAT expression by nearly 3-fold in COS-7 cells (Figure 1B). A 2-fold increase in CAT activity was also observed with a fragment of the DRR (fragment -2360 to -878 in construct p2BstE). These results indicate that the PRR contains sequences that repress the cognate promoter but are unable to inhibit the transcription of the heterologous TATA-containing tk promoter. In addition, the PRR bears elements capable of activating the heterologous promoter but not the *PSG-5* promoter. The lack of repression on the TATA-containing tk promoter might be explained by a requirement for a specific interaction of 5' upstream regulatory elements with basal transcription factors assembled on the *PSG-5* promoter.

Functional dissection of the PRR reveals multiple negative and positive regulatory elements

Regions analogous to the PRR of *PSG-5* have been reported to down-regulate the promoter activity of three other *PSG* genes in immortalized cells from human placenta [18,19]. However, they have not been further investigated. We therefore performed a detailed analysis of the PRR to obtain more knowledge about the organization and activity of this important region. Functional dissection of the PRR revealed an array of not only negative elements but also positive ones. The 166 bp fragment located just upstream of the promoter contains sequences that produced an almost 60% repression in HeLa cells and had essentially no effect in COS-7 cells in comparison with the control UBCAT plasmid (Figure 2). The inclusion of 5' sequences up to -457 and -524 successively increased the UB promoter activity in both cells. Construct -567CAT showed a CAT activity higher than that of UBCAT but lower than that of the -524CAT construct, mainly in COS-7 cells, in which promoter activation decreased

from 80% to only 5%, indicating that the -567 to -524 segment harbours a negative regulatory activity. Finally, the inclusion of the last 38 bp of the PRR caused a marked repression of CAT activity in both cell lines. Taken together, these experiments delineated the presence of activating elements located between positions -524 and -421, and two negative functions located between positions -605 and -524, which effectively operate in both cell types. In addition, our analysis revealed the presence of an inhibitory element located at -421 to -254 that acts predominantly in HeLa cells.

Identification of specific protein-DNA interactions within the PRR

Having identified different *cis* elements within the PRR, we performed footprinting assays *in vitro* to obtain a correlation between the functional observations and DNA-protein interactions. Two DNA fragments containing PRRa (-421 to -254) and PRRb (-605 to -422) were end-labelled and subjected to digestion with DNase I in the presence or absence of nuclear extracts from human term placenta and HeLa cells. These experiments revealed four protected regions of the PRRb probe with both extracts. Footprint 1 (FP1), FP2 and FP3 mapped within the functional activating segments (coordinates -455 to -433, -502 to -490 and -512 to -505 respectively) (Figure 3). Meanwhile, FP4 (-537 to -523) mapped within the inhibitory region described above. Within PRRa, two less evident protected sites were observed in the presence of placental and HeLa extracts (coordinates -312 to -293 and -285 to -271) (results not shown).

A sequence comparison of the FP1-FP4 footprint regions with similar regions of other members of the *PSG* gene family revealed that these sequences are highly conserved. In contrast, no significant similarities within FP1-FP4 were found in the corresponding 5' regulatory regions of the *CEA* subgroup of the human *CEA/PSG* gene family, i.e. the *CEA*, *NCA* and *BGP* genes, which are expressed in other cell types but not in placenta ([20,35,36], and results not shown). In addition, a close relative of *PSG*, the *CGM6* gene, which is expressed exclusively in granulocytes [37], showed high similarity only within FP3. These observations suggest that FP1, FP2 and FP4 could be involved

Figure 3 DNase I footprinting analysis of the *PSG-5* gene 5'-proximal regulatory region

The end-labelled DNA probe containing PRRb (–422 to –605) was incubated with nuclear extracts from human term placenta (25 and 50 μ g) and HeLa cells (15 and 30 μ g) and then subjected to partial digestion with DNase I as described in the Experimental section. Lane G + A represents the Maxam–Gilbert sequencing reaction [53]. The boundaries of the FP1, FP2, FP3 and FP4 sites are on the borders; the sequences of the protected regions are indicated below.

in regulating the tissue-specific expression of the *PSG* gene family members.

The specificity of the DNA–protein interactions at FP1 was further analysed through EMSA studies performed with protein extracts from human term placenta, HeLa and COS-7 cells. The incubation of placental nuclear extracts with the FP1 probe yielded a specific DNA–protein complex (P1) that was abolished by an excess of unlabelled FP1 competitor DNA but not by the same excess of an unrelated oligonucleotide (Figure 4, lanes 2–4 and 9–10). Two oligonucleotides containing internal mutations were employed as competitor probes to localize nucleotide sequences serving as binding sites for the assembly of the specific complex. The mutA oligonucleotide [FP1 A-box (GAGGGC) $\rightarrow$ mutA (ATCTAG)] was unable to compete for binding to the native sequence (Figure 4, lanes 5 and 6), whereas the mutB oligonucleotide [FP1 B-box (CGTGAA) $\rightarrow$ mutB (ATCAGC)] effectively displaced the complex (Figure 4, lanes 7 and 8). Thus the FP1 sequence is recognized as the binding site for at least one placental nuclear protein, and the A-box is necessary for complex assembly. This was confirmed by the inability of the mutA probe to form the complex, whereas the mutB probe rendered the same complex as the wild-type probe (results not shown).

Figure 4 EMSA of FP1-binding-site probe with cell extracts from human term placenta

Nuclear extracts from human term placenta were incubated with a 32 P-labelled FP1-site oligonucleotide in the absence or presence of unlabelled competitor FP1 site, mutated FP1 site (MutA and MutB) or non-specific (NS1) oligonucleotides. After incubation, reaction mixtures were subjected to gel electrophoresis and autoradiography. f, mobility of free probe.

FP1 oligonucleotide was also recognized by both HeLa and COS-7 cells extracts. HeLa extracts rendered three specific retarded bands (H1, H2 and H3) that were completely displaced by the FP1 competitor DNA (Figure 5A, lanes 2 and 3). A 100-fold molar excess of the mutA oligonucleotide effectively displaced radiolabelled probe from complex H1 and to a smaller extent from H2 but enhanced the formation of complex H3 (Figure 5A, lane 4). In contrast, a 100-fold molar excess of mutB oligonucleotide effectively inhibited the assembly of complexes H2 and H3 and did not interfere in the formation of complex H1 (Figure 5A, lane 5). Only complexes H1 and H2 were detected with the mutA oligonucleotide probe (Figure 5A, lanes 6–8). In contrast, the mutB probe allowed the assembly of complexes H2 and H3 but not H1 (Figure 5A, lanes 9–11). Taken together, these results demonstrate that the A-box is the binding site for proteins present in the major H3 complex, the B-box is necessary for the assembly of complex H1, and the formation of complex H2 requires sequences other than those mutated in oligonucleotides mutA and mutB.

The same set of experiments was performed simultaneously with COS-7 cell extracts. In this case, one abundant (C1) and three minor specific DNA–protein complexes were observed (Figure 5B). Oligonucleotide mutA effectively competed for all four complexes, whereas mutB clearly competed for the assembly of complexes C2 and C4 and, to a smaller extent, for C1 and C3 (Figure 5B, lanes 4 and 5). The four complexes were observed when the mutA oligonucleotide was used as probe, whereas the mutB probe allowed the assembly of all but the C1 complex (Figure 5B, lanes 6–9). Thus, in COS-7 cells, DNA-binding proteins interact with FP1 mainly through the B-box sequence and lastly through a sequence different from the B-box and the A-box.

In sum, DNA-binding proteins present in placental, HeLa and COS-7 cells recognize FP1 as a specific DNA-binding site. Sequence-dependent binding and electrophoretic patterns strongly suggest that several different DNA-binding proteins are involved in complex formation. Nevertheless, firm conclusions must await the identification and characterization of these DNA-binding proteins.

DNA–protein target sites present in the PRR are involved in *PSG-5* gene transcription

To investigate the functional relevance of the DNA–protein interactions defined in PRR, a mutation analysis was performed. Site-directed mutagenesis in the A-box or B-box of FP1 resulted in a further repression of CAT activity in HeLa, COS-7 and HTR8/SVneo cells in comparison with that of the PBCAT plasmid (Figure 6, PBmutA and PBmutB). These results indicate that FP1 is the binding site for activating factors required to maintain promoter activity. When FP3 was mutated in the context of the whole PRR, a slightly higher repression of the promoter was observed in HeLa and COS-7 cells (Figure 6, PBmFP3), suggesting that it also corresponded to the binding site of an activating factor. However, the simultaneous mutation of FP3 and deletion of the 80 bp repressor region clustered at –605 to –524 resulted in nearly 300% activation in comparison with the control UBCAT plasmid and a 2-fold increase over the parental –524CAT construct. Increased CAT activity of this construct was also observed in HTR8/SVneo placental cells (Figure 6, –524CATmFP3). We interpret these results as a repression of *PSG-5* transcription by DNA-binding proteins bound at the FP3 site acting in concert with other factors bound within the upstream 80 bp segment. Disruption of the negative elements permits the activation of the promoter by enhancers, probably those located at FP1 and FP2. Importantly, this activation effectively operates in placental and in non-placental cell types.

DISCUSSION

The *PSG-5* gene belongs to the *PSG* branch of the *CEA* gene family. Expression of the *PSG* subgroup is highly restricted to developing syncytiotrophoblast cells and to a few other cell types [8–12]. We have previously reported the characterization of an essential element of *PSG-5* promoter, CPE, that contributes to its basal activity in different cell types [15]. To obtain further insight into the regulatory events implicated in *PSG* gene expression, we have analysed an extensive regulatory region of the *PSG-5* gene. We have demonstrated the existence of positive and negative *cis*-acting elements responsible for similar transcription patterns in non-placental HeLa and COS-7 cells, which respectively do and do not express *PSG* genes. In both cell types, *PSG-5* promoter (–254 to –43) is active but down-regulated by at least two upstream promoter-selective negative regulatory regions, the PRR and DRR modules. In addition, within the 3 kb 5' upstream sequence and the first intron there are apparently no enhancers that are able to relieve repression. More relevantly, we have demonstrated that the PRR and DRR also behave as repressor elements in the placental cell line HTR8/SVneo and that the UB fragment of *PSG-5* exhibits essentially the same promoter activity in the three cell types employed in our study. These results indicate clearly that the –3204 to –43 sequence of the *PSG-5* gene does not itself convey tissue-specific expression. However, we cannot rule out the possibility that it harbours regulatory elements that respond in a tissue-specific manner to a developmental or extracellular stimulus.

The negative regulatory activity of the PRR required an intact FP3 motif (–512 to –505) and the fragment –605 and –524, which bears the FP4 site (–537 to –523). FP3-binding and FP4-binding activities were detected in nuclear extracts obtained from cells in which *PSG* genes are transcribed, suggesting that they are not involved in cell-specific repression but they might mediate gene activation in a developmental and/or cell-specific manner. The search for consensus sequences in a library of transcription-

factor-binding sites [38] revealed no relevant similarity of FP3 and FP4 to known regulatory motifs, indicating that they might be the target site for novel transcription factors.

Positive regulatory elements were present between positions –524 and –421. This segment included FP2-binding and FP1-binding sites. The FP2 sequence showed no significant similarity to known *cis*-acting motifs; its DNA-binding proteins await further characterization. The FP1 protein-binding site formed multiple DNA–protein complexes with placenta, HeLa and COS-7 cell extracts. Although the DNA–protein binding patterns were somewhat different, the FP1 sequence behaved as a positive transcriptional control element in the three cell lines tested, sustaining promoter activity. Several high-score putative binding sites for FP1 were found, such as a perfect match to the binding site for the gut Krüppel-like factor ('GKLF') (G/AG/AGGC/TGC/T) [39], a putative hepatic nuclear factor 4 (HNF-4) binding sequence and a peroxisome-proliferator-responsive element ('PPRE'). The comparison of the FP1 sequence with the consensus sequences required for the binding of the HNF-4 transcription factor (NNNRGGNCAAAGK-TCANN) and the PPAR-RXR (peroxisome-proliferator-activated receptor-retinoid X receptor) heterodimer (CWRAW-CTAGGNCAAAGGTCA) [40] shows the coincidence of 11 of 13 and 15 of 20 bases respectively. Wendling et al. [41] have recently established that RXRs, through heterodimerization with RARs and probably with PPAR γ , are essential for embryonic development and placentogenesis. Therefore these nuclear receptors might be also involved in controlling placenta-specific genes such as *PSGs*.

The sequences of the FP1–FP4 regions are highly conserved in all known members of the *PSG* gene family subgroup expressed in the syncytiotrophoblast. In contrast, FP1–FP4 sequences diverge in genes of the *CEA* subgroup and a significant similarity to the FP3 is found only in the *CMG6* gene, which is expressed exclusively in granulocytes [37]. Conservation of the FP1, FP2 and FP4 sequences within the *PSG* genes but not the *CEA* subgroup genes indicates that they are probably important for mediating the tissue-specific transcriptional activation of *PSG* genes. Certainly, a detailed mutational and functional analysis of the footprint sequences would be required for a better understanding of this issue.

So far, most of the silencers described have exerted their repressive function in a promoter-independent fashion (reviewed in [34]). Nevertheless, some promoter-dependent silencers have been reported, such as the rep27 silencer element in the rat *PAP I* (pancreatitis-associated protein I) gene [42]. This type of silencer might decrease the frequency of transcription initiation via a promoter-specific factor, such as a promoter-specific TATA-binding-protein associated factor ('TAF') or mediator component [34]. The *PSG-5* gene promoter, as well as all *PSG* family members, lacks a TATA-box, large GC-regions or a typical pyrimidine-rich initiator. This special promoter context further emphasizes that the mechanism for repression mediated by the PRR and DRR is due to interactions with specific *PSG-5* promoter elements. Our observations are consistent with the proposal that the differences between TATA-containing and TATA-less promoters might allow them to respond to a different subset of activators and/or repressors [43].

The promoter activity of other *PSG* family members, i.e. *PSG1-I*, *PSG12* and *PSG6* genes assayed in placental HP-A1 cells transformed by simian virus 40 [44], was also inhibited by upstream sequences up to positions –834, –835 and –614 respectively [18,19]. Taken together, these observations and our present results strongly suggest that essentially the same regulatory events take place within the proximal 5' regulatory regions

Figure 7 Hypothetical model of *PSG-5* gene transcription

The putative regulatory elements including positive FP1 and FP2, the negative FP3, FP4, PRRa (−421 to −254) and the DRR (−3204 to −706) regulatory regions identified in this study are indicated. Upper panel: repressed state, in which repressors down-regulate *PSG-5* promoter and activators are unable to override such a repression. Lower panel: activated state, in which repression activity is impaired and activators are able to enhance the transcription of *PSG-5* promoter.

of *PSG* genes in placental and non-placental cell types. Therefore *PSG* genes are tightly repressed under the experimental conditions employed in these analyses. Therefore how can *PSG* transcriptional activation, especially the increase during placental development, be explained? Several possibilities can be postulated. First, the activation of *PSG* genes could be mediated by *cis* elements located farther 5' upstream, downstream, or even within the *PSG* gene family cluster. Tissue-specific transcriptional activation relying on these types of enhancer sequence has been demonstrated for many genes, such as the human placental lactogen gene whose expression depends on a tissue-specific enhancer found downstream of the three human genes encoding placental lactogen [45,46]. Secondly, a locus control region with a very strong enhancer activity might function in the formation of open or active chromatin structure extending over the entire *PSG* locus. Examples of genes regulated by such elements include the human β -globin, the chicken lysozyme loci and the tyrosinase gene [47–49]. Thirdly, the tissue-specific and development-specific expression of *PSG* genes might be controlled by a derepression mechanism. Several studies have strongly suggested the importance of such a mechanism controlling the specific activation of gene expression. For example, derepression has been postulated in the induction of the gene for prodynorphin by protein kinase A, in the cell cycle regulation of the cyclin A promoter and in the expression of the gene for C/EBP (CCAAT-enhancer-binding protein) during adipogenesis [50–52].

The results presented here support the following working model to explain the regulatory events that take place within the 5' region of the *PSG-5* gene (Figure 7). The singular characteristics of the *PSG-5* promoter determine the assembly of a preinitiation complex that is selectively down-regulated by repressors bound at FP3, FP4 and probably between −605 and −567. Two additional negative regulatory elements, DRR and PRRa (−421 to −254), collaborate to tightly down-regulate *PSG-5* gene expression. Proteins that recognize FP1 and probably

FP2 as their target-binding sites are devoid of activating *PSG-5* promoter in the presence of the functional repressor complexes. Transcriptional activation of the *PSG-5* gene would be achieved through an impairment of the repressing functions. This could be due to the interplay of the regulatory elements identified in this paper with intermediary positive factors that are absent or inactive at the initial onset of placentation and under culture conditions *in vitro*. Alternatively, silencer activity might be switched off simply by a decrease in the amount or the activity of negative regulatory factors during gestation.

Finally, further investigations of extracellular stimuli that could act as specific inducers or derepressors might help to ascertain the physiological regulation of *PSG* genes.

We thank Dr C. H. Graham for providing us with the HTR8/SVneo cells; Dr Susana Genti-Raimondi and Daniela Slavin for many helpful discussions; and Dr Ana Cabanillas for critical reading of the manuscript. This work was supported by grants from the Consejo Nacional de Investigaciones Científicas y Tecnológicas de Argentina (CONICET), the Agencia Nacional de Promoción Ciencia y Técnica, the Consejo de Investigaciones Científicas y Tecnológicas de la Provincia de Córdoba and the Secretaría de Ciencia y Técnica de la Universidad Nacional de Córdoba. J.L.B. and N.P.K. are Investigators of CONICET. C.I.D. thanks CONICET for the fellowship.

REFERENCES

- 1 Cross, J. C., Werb, Z. and Fisher, S. J. (1994) Implantation and the placenta: key pieces of the development puzzle. *Science* **266**, 1508–1518.
- 2 Tatarinov, Y. S. and Masyukevich, V. N. (1970) Immunological identification of a new β 1-globulin in the blood serum of pregnant women. *Byull. Eksp. Biol. Med. USSR* **69**, 66–68.
- 3 Wurz, H., Geiger, W., Kunzig, H. J., Jabs-Lehmann, A., Bohn, H. and Luben, G. (1981) Radioimmunoassay of SP1 (pregnancy-specific beta1-glycoprotein) in maternal blood and in amniotic fluid normal and pathologic pregnancies. *J. Perinat. Med.* **9**, 67–78.
- 4 Tamsen, L., Johansson, S. G. and Axelsson, O. (1983) Pregnancy-specific beta 1-glycoprotein (SP1) in serum from women with pregnancies complicated by intrauterine growth retardation. *J. Perinat. Med.* **11**, 19–25.
- 5 Khan, W. N., Teglund, S., Brener, K. and Hammarström, S. (1992) The Pregnancy-Specific Glycoprotein family of the immunoglobulin superfamily: identification of new members and estimation of family size. *Genomics* **12**, 780–787.
- 6 Olsen, A., Teglund, S., Nelson, D., Gordon, L., Copeland, A., Georgescu, A., Carrano, A. and Hammarström, S. (1994) Gene organization of the pregnancy-specific glycoprotein region on human chromosome 19: assembly and analysis of a 700-kb cosmid contig spanning the region. *Genomics* **23**, 659–668.
- 7 Hammarström, S. (1999) The carcinoembryonic antigen (CEA) family: structures, suggested functions and expression in normal and malignant tissues. *Semin. Cancer Biol.* **9**, 67–81.
- 8 Bocco, J. L., Panzetta, G., Flury, A. and Patrito, L. C. (1989) Expression of pregnancy-specific β 1-glycoprotein gene in human placenta and hydatidiform mole. *Biochem. Int.* **18**, 999–1008.
- 9 Chan, W. Y., Tease, L. A., Borjigin, J., Chan, P. K., Rennert, O. M., Srinivasan, B., Shupert, W. L. and Cook, R. G. (1988) Pregnancy-specific β 1-glycoprotein mRNA is present in placental as well as non-placental tissues. *Hum. Reprod.* **3**, 677–685.
- 10 Wu, S. M., Bazar, L. S., Cohn, M. L., Cahill, R. A. and Chan, W.-Y. (1993) Expression of pregnancy-specific β 1-glycoprotein genes in hematopoietic cells. *Mol. Cell. Biochem.* **122**, 147–158.
- 11 Plouzek, C. A., Leslie, K. K., Stephens, J. K. and Chou, J. Y. (1993) Differential gene expression in the amnion, chorion and trophoblast of the human placenta. *Placenta* **14**, 277–285.
- 12 Leslie, K. K., Watanabe, S., Lei, K.-J., Chou, D. Y., Plouzek, C. A., Deng, H. C., Torres, J. and Chou, J. Y. (1990) Linkage of two human pregnancy-specific β 1 glycoprotein genes: one is associated with hydatidiform mole. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 5822–5826.
- 13 Panzetta-Dutari, G. M., Bocco, J. L., Reimund, B., Flury, A. and Patrito, L. C. (1992) Nucleotide sequence of a pregnancy-specific β 1 glycoprotein gene family member. Identification of a functional promoter region and several putative regulatory sequences. *Mol. Biol. Rep.* **16**, 255–262.
- 14 Thompson, J., Koumari, R., Wagner, K., Barnert, S., Schleussner, C., Schrewe, H., Zimmermann, W., Müller, G., Schempp, W., Zaninetta, D. et al. (1990) The human pregnancy-specific β 1-glycoprotein genes are tightly linked on the long arm of chromosome 19 and are coordinately expressed. *Biochem. Biophys. Res. Commun.* **167**, 848–859.

- 15 Koritschoner, N. P., Panzetta-Dutari, G. M., Bocco, J. L., Dumur, C. I., Flury, A. and Patrito, L. C. (1996) Analyses of *cis*-acting and *trans*-acting elements that are crucial to sustain pregnancy-specific glycoprotein gene expression in different cell types. *Eur. J. Biochem.* **236**, 365–372
- 16 Koritschoner, N. P., Bocco, J. L., Panzetta-Dutari, G. M., Dumur, C. I., Flury, A. and Patrito, L. C. (1997) A novel human zinc finger protein that interacts with the core promoter element of a TATA box-less gene. *J. Biol. Chem.* **272**, 9573–9580
- 17 Slavin, D., Sapin, V., López-Díaz, F., Jacquemin, P., Koritschoner, N., Dastugue, B., Davidson, I., Chatton, B. and Bocco, J. L. (1999) The Kruppel-like core promoter binding protein gene is primarily expressed in placenta during mouse development. *Biol. Reprod.* **61**, 1586–1591
- 18 Lei, K.-J., Wang, C., Chamberlin, M. E., Liu, J. L., Pan, C. J. and Chou, J. Y. (1993) Characterization of two allelic variants of a human pregnancy-specific glycoprotein gene. *J. Biol. Chem.* **268**, 17528–17538
- 19 Chamberlin, M. E., Lei, K.-J. and Chou, J. Y. (1994) Subtle differences in human pregnancy-specific glycoprotein gene promoters allow for differential expression. *J. Biol. Chem.* **269**, 17152–17159
- 20 Hauck, W., Nédellec, P., Turbide, C., Stanners, C. P., Barnett, T. R. and Beauchemin, N. (1994) Transcriptional control of the human biliary glycoprotein gene, a CEA gene family member down-regulated in colorectal carcinomas. *Eur. J. Biochem.* **223**, 529–541
- 21 Keck, U., Nédellec, P., Beauchemin, N., Thompson, J. and Zimmermann, W. (1995) The cea10 gene encodes a secreted member of the murine carcinoembryonic antigen family and is expressed in the placenta, gastrointestinal tract and bone marrow. *Eur. J. Biochem.* **229**, 455–464
- 22 Strydio, C., Swillens, S., Georges, M., Szpirer, C. and Vassart, G. (1990) Structure, evolution and chromosomal localization of the human pregnancy-specific beta 1 glycoprotein gene family. *Genomics* **6**, 579–592
- 23 Luckow, B. and Schütz, G. (1987) CAT construction with multiple unique restriction sites for the functional analysis of eukaryotic promoter and regulatory elements. *Nucleic Acids Res.* **13**, 5490–5491
- 24 Sanger, F., Nicklen, N. and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
- 25 Banerji, J. S., Rusconi, S. and Schaffner, W. (1981) Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**, 299–308
- 26 Gorman, C. M., Moffat, L. F. and Howard, B. H. (1982) Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* **2**, 1044–1051
- 27 Dignam, J. D., Lebowitz, R. M. and Roeder, R. G. (1983) Accurate transcription initiation by polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**, 1475–1489
- 28 Finocchiaro, L. M. E., Amati, P. and Gilkin, G. C. (1992) Single strand binding protein specific for the polyoma early-coding strand of PEA-1 (AP-1) regulatory sequence. *Nucleic Acids Res.* **19**, 4279–4287
- 29 Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254
- 30 Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (1994) *Current Protocols in Molecular Biology*, Green Publishing Associates and J. Wiley & Sons, New York
- 31 Cereghini, S., Raymondjean, M., Garcia Carranca, A., Herbomel, P. and Yaniv, M. (1987) Factors involved in control of tissue-specific expression of albumin gene. *Cell* **50**, 627–638
- 32 Reference deleted
- 33 Graham, C. H., Hawley, T. S., Hawley, R. G., MacDougall, J. R., Kerbel, R. S., Khoo, N. and Lala, P. K. (1993) Establishment and characterization of first trimester human trophoblast cells with extended lifespan. *Exp. Cell Res.* **206**, 204–211
- 34 Ogbourne, S. and Antalís, T. M. (1998) Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes. *Biochem. J.* **321**, 1–14
- 35 Schrewe, H., Thompson, J., Bona, M., Hefta, L. J., Maruya, A., Hassauer, M., Shively, J. E., von Kleist, S. and Zimmermann, W. (1990) Cloning of the complete gene for carcinoembryonic antigen: analysis of its promoter indicates a region conveying cell type-specific expression. *Mol. Cell. Biol.* **10**, 2738–2748
- 36 Oikawa, S., Kosaki, G. and Nakazato, H. (1987) Molecular cloning of a gene for a member of carcinoembryonic antigen (CEA) gene family; signal peptide and N-terminal domain sequences of nonspecific crossreacting antigen (NCA). *Biochem. Biophys. Res. Commun.* **146**, 464–469
- 37 Eades-Perner, A. M., Thompson, J., van der Putten, H. and Zimmermann, W. (1998) Mice transgenic for the human CGM6 gene express its product, the granulocyte marker CD66b, exclusively in granulocytes. *Blood* **91**, 663–672
- 38 Quandt, K., Frech, K., Karas, H., Wingender, E. and Werner, T. (1995) MatInd and MatInspector, new fast and versatile tools for the detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res.* **23**, 4878–4884
- 39 Shields, J. M. and Yang, V. W. (1998) Identification of the DNA sequence that interacts with the gut-enriched Kruppel-like factor. *Nucleic Acids Res.* **26**, 796–802
- 40 Palmer, C. N. A., Hsu, M.-H., Griffin, K. J. and Johnson, E. F. (1995) Novel sequence determinants in peroxisome proliferator signaling. *J. Biol. Chem.* **270**, 16114–16121
- 41 Wendling, O., Chambon, P. and Mark, M. (1999) Retinoid X receptors are essential for early mouse development and placentogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 547–551
- 42 Ortiz, E. M., Dusetti, N. J., Dagorn, J. C. and Iovanna, J. L. (1997) Characterization of a silencer regulatory element in the rat PAP I gene which confers tissue-specific expression and is promoter-dependent. *Arch. Biochem. Biophys.* **340**, 111–116
- 43 Smale, S. T. (1997) Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. *Biochim. Biophys. Acta* **1351**, 73–88
- 44 Lei, K.-J., Gluzman, Y., Pan, C.-J. and Chou, J. Y. (1992) Immortalization of virus-free human placental cells that express tissue-specific functions. *Mol. Endocrinol.* **6**, 703–712
- 45 Jacquemin, P., Oury, C., Peers, B., Morin, A., Belayew, A. and Martial, J. A. (1994) Characterization of a single strong tissue-specific enhancer downstream from the three human genes encoding placental lactogen. *Mol. Cell Biol.* **14**, 93–103
- 46 Walker, W. H., Fitzpatrick, S. L., Barrera-Saldana, H. A., Resendez-Perez, D. and Saunders, G. F. (1991) The human placental lactogen genes: structure, function, evolution and transcriptional regulation. *Endocr. Rev.* **12**, 316–328
- 47 Stamatoyannopoulos, G. and Neinhuis, A. W. (1994) Hemoglobin switching. in *The Molecular Basis of Blood Diseases* (Stamatoyannopoulos, G., Neinhuis, A. W., Majerus, P. and Varmus, H., eds), pp. 107–155, W. B. Saunders, Philadelphia
- 48 Bonifer, C., Vidal, M., Grosfeld, F. and Sippel, A. E. (1990) Tissue specific and position independent expression of the complete gene domain for chicken lysozyme in transgenic mice. *EMBO J.* **9**, 2843–2848
- 49 Montoliu, L., Umland, T. and Schütz, G. (1996) A locus control region at —12 kb of the tyrosinase gene. *EMBO J.* **14**, 6026–6034
- 50 Carrión, A. M., Mellström, B. and Naranjo, J. R. (1998) Protein kinase A-dependent derepression of the human prodynorphin gene via differential binding to an intragenic silencer element. *Mol. Cell. Biol.* **18**, 6921–6929
- 51 Liu, N., Lucibello, F. C., Engeland, K. and Muller, R. (1998) A new model of cell cycle-regulated transcription: repression of the cyclin A promoter by CDF-1 and anti-repression by E2F. *Oncogene* **16**, 2957–2963
- 52 Jiang, M.-S., Tang, Q.-Q., McLennan, J., Geiman, D., Shillinglaw, W., Henzel, W. J. and Lane, M. D. (1998) Derepression of the C/EBP α gene during adipogenesis: identification of AP-2 α as a repressor. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 3467–3471
- 53 Maxam, A. M. and Gilbert, W. (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**, 499–540

Received 22 November 1999/18 May 2000; accepted 23 June 2000