Universiteit Leiden
The Netherlands

Regulation of human protein S gene (PROS1) transcription

Wolf, Cornelia de

Citation
Wolf, C. de. (2006, May 29). *Regulation of human protein S gene (PROS1) transcription*. Retrieved from https://hdl.handle.net/1887/4413

Version: Corrected Publisher's Version

License: Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden

Downloaded from: https://hdl.handle.net/1887/4413

Note: To cite this publication please use the final published version (if applicable).

Chapter 2

Initiation of Protein S mRNA synthesis in human liver, various cell lines, and Protein S promoter-reporter gene Plasmids

Part of this chapter was published as

Initiation of Protein S mRNA synthesis in
human liver and various cell lines.

De Wolf CJ, Cupers RM, Bertina RM, and Vos HL.

J Thromb Haemost. (2005) 3(2):410-412

Summary

Protein S (PS), a cofactor of activated protein C (APC), is a vitamin K-dependent anticoagulant plasma protein, which is produced mainly in the liver but which is also produced extrahepatically. Therefore, the production of this protein may be regulated in a tissue-specific manner. Upon analysis we found that the 5'-flanking region of the human PS gene (*PROS1*) lacks characteristic "CCAAT" and "TATA" boxes. To determine the sites of transcriptional initiation *PROS1* transcripts from HepG2, HuH7, HeLa, HUVEC, and Meg01 cells and human liver were subjected to transcription start site (TSS) analysis by using 5'-rapid amplification of cDNA ends. In all cell types, as well as in human liver, transcription is initiated most frequently at one of three TSSs located 100 bp, 117 bp or 147 bp upstream from the translational startcodon. HUVEC cells contained an additional TSS at –200 bp. In all cell lines, reporter constructs containing a minimal promoter of 370 bp upstream of the translational startcodon demonstrated maximal promoter activity. Whereas three distinct TSSs were identified for the endogenous *PROS1* transcripts, no preferred TSSs could be determined for transiently transfected *PROS1* promoter luciferase constructs. This is consistent with our finding that deletion of the first 181 bp proximal to the translational start in these constructs did not result in complete deterioration of promoter activity. This suggests a relaxation of TSS regulation when using plasmid constructs in transient transfection studies.

Introduction

Reduced plasma Protein S (PS) levels were first reported to be involved in the development of venous thrombo-embolism (VTE) in 1984 (1-3). PS deficiency has been well established as a risk factor for VTE since then (4-7). PS is a vitamin K-dependent plasma protein that displays anticoagulant properties by acting as a non-enzymatic cofactor for activated Protein C (APC) in the proteolytic degradation of blood clotting factors Va and VIIIa (8-11). In human plasma, PS circulates in an active free form (40%) and a C4b-binding protein (C4BP)-bound inactive form (60%) (12-14). The major source of circulating plasma PS is the hepatocyte (15). However, PS is also known to be produced by a variety of other cell types such as megakaryocytes (16;17), endothelial cells (18;19), Leydig cells (20), osteoblasts (21) and cells of the nervous system (22). In mammals, PS mRNA was found in virtually all tissues and organs examined (23).

Two copies of the PS gene are located on chromosome 3. The active PS gene (*PROS1*) shares 96% homology with the inactive pseudogene (*PROS2*). The pseudogene, which lacks the promoter and the first exon also contains several frame-shift deletions leading to premature stopcodons that render this gene inactive (24-26). Mutational analysis of the exons of *PROS1* in patients with PS deficiency has shown various mutations within the PS gene to be responsible for low PS levels in plasma [reviewed by Gandrille *et al* (27;28)]. In contrast to the coding regions of the *PROS1* gene, the promoter region has been poorly investigated. As with other promoter regions from genes coding for vitamin K-dependent coagulation proteins, a distinct TATA-box is absent from the *PROS1* promoter region (29;30). The literature to date describes one primer extension assay (25) and multiple cDNAs derived from cDNA libraries (31-33). The databases at the National Center for Biotechnology Information (NCBI) provide multiple direct submissions of *PROS1* cDNAs (34). However, none of these cDNAs are necessarily full-length, whereas primer extension studies are notoriously difficult to interpret.

In the present study, we describe an analysis of the transcriptional control region of *PROS1*. To our knowledge for the first time we provide a transcription start site (TSS) distribution after analysis of multiple full-length *PROS1* cDNAs from various cell lines and human liver. Using a method that targets capped mRNA for reverse transcription we show

that there are a number of major start sites in the promoter of the endogenous *PROS1* gene. The cell types that were tested displayed small differences in TSS usage. In addition, we show that this preference for specific positions is lost in luciferase reporter constructs. The implications of these findings are discussed.

Experimental procedures

Plasmids --- The *PROS1* promoter reporter constructs used in this study, originated from a 7 kb *Eco*RI promoter fragment which was isolated from BAC clone #2513H18 from the CITBI-E1 genomic library (Research Genetics, Invitrogen, Carlsbad, CA) and ligated into the *Eco*RI site in the multiple cloning site region (MCS) of the pcDNA3 cloning vector (Invitrogen). This pcDNA3-*PROS1* construct contained 716 bp of sequence downstream from the translational startcodon. The complete sequence was determined through automated sequencing (ABI PRISM and Beckman CEQ2000 sequencers) and deposited in Genbank under accession number AY605182. Subsequently, pcDNA3- *PROS1* was modified to contain nucleotides -5948 to -1 from the translational startcodon as follows. First, the construct was digested with *Bam*HI and *Eco*RI (all restriction enzymes were obtained from New England Biolabs, Hertfordshire, UK) resulting in two fragments. Fragment -5948/-410 contained a pcDNA3-*Bam*HI and a *PROS1*-*Bam*HI terminus, and fragment -410/+716 contained a *PROS1*-*Bam*HI and a pcDNA3-*Eco*RI terminus. Fragment -410/+716 was ligated into pcDNA3, which had been linearized by digestion with *Bam*HI and *Eco*RI. The pcDNA3-*PROS1*-410/+716 was used to generate a product spanning from bp -410 to -1 by polymerase chain reaction (PCR) using the T7 primer in pcDNA3 and a mutant reverse gene-specific primer (GSP), PSstart, in which the initiation methionine was modified into an *Eco*RI site (Table 1). After digestion with *Bam*HI and *Eco*RI this PCR product was ligated into pcDNA3, which had been linearized by digestion with the same enzymes. *PROS1* fragment -5948/-410 was then religated into the *Bam*HI site connecting the 5' terminus of *PROS1*-410/-1 and pcDNA3. The resulting *PROS1* fragment (-5948/-1) was cloned directly 5' to the luciferase reporter gene in the pGL3basic vector (pGL3b, Promega, Madison, WI) after digestion with MCS restriction sites *Kpn*I and *Xho*I. This construct was named PS5948. PS5948 was linearized with *Kpn*I and *Nde*I (located at position –5798) and subsequently subjected to exonuclease III digestion (Erase a Base kit, Promega). The size of the resulting 5' deletion was determined by sequence analysis. The 5'-deletion constructs were used for transient transfection assays. After digestion of construct PS1062 with *Pst*I and *Eco*RI, exonuclease III

digestion was also used to obtain 3'-deletion fragments from this construct. Additional 3'-deletion constructs, PS-370/-181, PS-1062/-256 and PS-370/-256, were created using restriction enzymes. PS-370/-181 was created by combination of constructs PS-370 and PS-1062/-181 after digestion with *Hin*dIII. PS-1062/-256 and PS-370/-256 originated from the original PS370 and PS1062 constructs by religation after excision of a *Stu*I-*Eco*RI PS fragment (-256/-1). An overview of all luciferase reporter constructs is given in Figure 1.

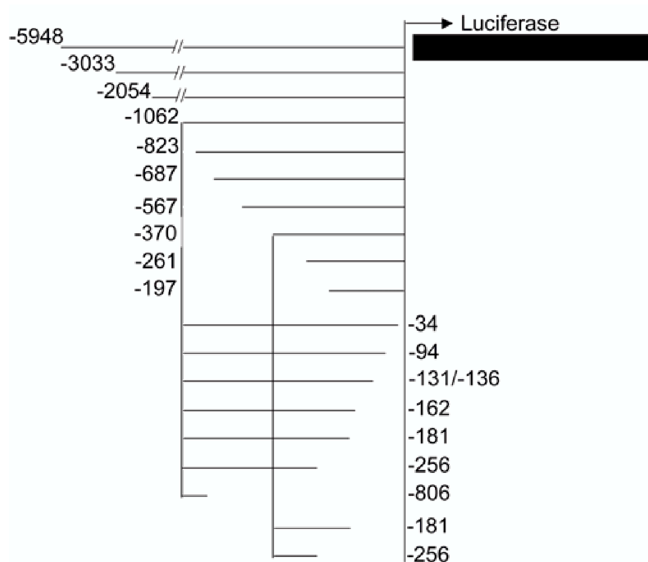


Figure 1 *PROS1* promoter-luciferase constructs in pGL3basic. The numbers indicate the length of the *PROS1* 5' flanking region proximal to the luciferase gene in pGL3basic. Constructs PS1062 and PS370 were also used in 3' deletion studies. The length of the 3' indentation is shown at the 3' end of the construct.

Human liver, cell lines and media --- A human liver sample was obtained from a deceased healthy donor and was a kind gift from E. de Wit at TNO Center for Prevention and Health, Leiden (TNO). The human hepatoblastoma cell line HepG2, and cervical adenocarcinoma cell line HeLa were purchased from the American Type Culture Collection (ATCC, Manassas, VA). Megakaryocytic cell line Meg01 was a kind gift from G. van Willigen from the Utrecht Medical Centre. The hepatoma cell line HuH7 was a kind gift from M. Verschuur (TNO). Primary human umbilical vein endothelial cells, HUVEC cells, were a kind gift from J. Grimbergen (TNO). HepG2, HuH7, HeLa cells were grown in Minimal essential medium

(MEM), 10% Fetal Bovine Serum (FBS), 100 μg/ml penicillin, 100 μg/ml streptomycin, and 1x MEM non-essential amino acids (all purchased from Gibco, Invitrogen). HUVECs were grown in M199-medium (BioWhittaker, Walkersville, MD), 10% heat inactivated human serum (local bloodbank), 10% Newborn Calf Serum (home made), 10 IE/ml heparin (BioWhittaker), 150 U/ml endothelial cell growth factor, 100 μg/ml penicillin (BioWhittaker) and 100 μg/ml streptomycin (BioWhittaker). 24 hrs before transfection the medium was replaced with heparin-free medium to prevent interference with the transfection assay. For PS antigen level measurements HUVEC cells were grown in serum-free EBM medium (BioWhittaker). Meg01 cells were grown in RPMI 1640 medium, 20% FBS, 100 μg/ml penicillin, 100 μg/ml streptomycin (all purchased from Gibco, Invitrogen). For PS antigen level measurements Meg01 cells were suspended at $1*10^6$ cells/ml in RPMI 1640 medium with 10% FBS.

Reporter gene assays --- $1*10^6$ Meg01 suspension cells were used per transfection. All adherent cell lines (HepG2, HuH7, HeLa, HUVEC) were transfected at 60-80% confluency. HUVEC cells were transfected in passage 2-3, whereas the other cell types were used up to passage 25. Each transfection was performed in triplicate in 12-wells plates, all assays were conducted with two different DNA preparations of each construct. Transfections with HepG2, HeLa, HUVEC and HuH7 cell lines were carried out using 3 μl Tfx-20 lipids (Promega) per μg transfected DNA. Meg01 cells were transfected using 10 μg DAC-30 (Eurogentec, Seraing, Belgium) per 2 μg DNA. In each transfection an equimolar concentration of construct was used supplemented with pUC13-MCS vector to obtain a fixed amount of transfected DNA. In pUC13-MCS the MCS had been removed by digestion with *Pvu*II and religation. Control vector pRL-SV40 (Promega), expressing the Renilla luciferase, was co-transfected for correction of transfection efficiency in a 1:500 ratio to total transfected DNA in HepG2, HuH7 and HeLa cell lines, a 1:100 ratio in transfections with HUVEC and Meg01 cells. The cell extracts were harvested at either 24 hours (HepG2, HuH7) or 48 hours (Meg01, HUVEC, HeLa) after transfection. Luciferase activity was measured according to the Dual Luciferase Assay System Protocol (Promega). All cell lines were lysed in 250 μl Passive Lysis Buffer/well, after which 20 μl was used to measure luciferase activity in HepG2 and HuH7 cells, 100 μl was used for HeLa, HUVEC and Meg01 cells. Activity was measured using a Lumat LB9507 luminometer (Berthold, Bad Wildbad, Germany).

PS measurements --- Total PS antigen levels in culture media were determined by enzyme-linked immunosorbent assay (ELISA) as described previously (35), with the following

modifications. ELISA plates were coated with goat anti-human PS IgG (Kordia, Leiden, The Netherlands) overnight at 4°C. A second coating with 2.5% ovalbumin (Sigma-Aldrich, St. Louis, MO) at 37°C for 1 hour was performed to reduce background absorbance. Complexes were detected with horseradish peroxidase-conjugated rabbit anti-human IgG (Dako, Glostrup, Denmark). Absorbance at 450 nm was determined with an Organon Teknika plate reader (Turnhout, Belgium).

RNA-assays --- Total RNA was isolated from cell culture or frozen tissue using Trizol reagent (Invitrogen) according to the manufacturers recommendations. Samples were treated with RNAse-free DNAse I (Amersham, Roosendaal, The Netherlands) after which RNA was purified with the RNeasy mini kit (Qiagen, Hilden, Germany). For 5' RACE (Rapid Amplification of cDNA Ends) experiments polyA RNA was isolated from total RNA preparations with a PolyA isolation kit (Ambion, Austin, TX). *PROS1* RNA levels were determined by quantitative real-time PCR analysis (QPCR). First, 1µg total RNA from each cell line was reverse transcribed using Superscript II reverse transcriptase and random hexamers (Invitrogen). 1/20th of the obtained cDNA was subsequently used in a QPCR reaction with primers and probes specific for *PROS1*. The primers and probe sequence locations and lengths were determined by using the ABI Primer Express Program (Applied Biosystems, Foster City, CA), and custom synthesized at Eurogentec. *PROS1* QPCR reactions (Eurogentec) were performed in 0.5 ml thin-walled, optical-grade PCR tubes (Applied Biosystems) in a 50 µl final volume, by addition of the following components: 0.25 U AmpliTaq Gold DNA polymerase, 160 nM TaqMan probe, 300 nM of each primer, and 3 mM $MgCl_2$. A QPCR of the internal standard, the porphobilinogen deaminase gene (PBGD), was carried out in a similar fashion for each RNA sample with 4 mM $MgCl_2$. An Applied Biosystems Prism model 7700 sequence detection instrument monitored the reactions. Thermal cycling conditions consisted of 10 min at 95°C followed by 50 cycles of 15 s at 95°C and 1 min at 60°C. Determinations of cycle threshold (C_T) were performed automatically by the instrument. The results are expressed as fold transcript relative to the internal standard PBGD ($=2^{\Delta Ct}$).

Determination of the TSS --- TSS analysis was performed on polyA RNA isolated from untransfected and transfected cell lines and human liver. The procedure was carried out using the Gene Racer™ kit (Invitrogen). The amount of polyA RNA used for the experiment varied from 300 to 2000 ng and was dependent on the amount of protein and mRNA detected in ELISA and QPCR analysis. After selection of full-length mRNA with the Gene Racer Kit,

first strand cDNA synthesis was performed with Superscript II reverse transcriptase and random hexamers (Invitrogen). For the untransfected cell lines and human liver the gene-specific amplification of full-length cDNA was performed with a 5'-primer provided in the Gene Racer kit and a 3' GSP located in the first exon of the *PROS1* mRNA (PSex1). For the cell lines transfected with a *PROS1* promoter reporter construct, the GL2 primer (Promega) was used as GSP. A *PROS1* primer in exon 2 was used for amplification of *PROS1* transcripts from a more downstream position (PSex2). The sequence and location of all non-commercially available primers are depicted in Table 1. The *PROS1* PCR cycle conditions were as follows; 5 cycles with an annealing temperature (T_a) of 72°C, 5 cycles with T_a 70°C, 25 cycles with T_a 68°C. After 20 cycles an equal volume of fresh PCR mix was added. For the construct PCR the annealing temperatures were, 68°C, 66°C, and 64°C respectively. The elongation time was set to 1.5 minutes to ensure complete elongation of all possible transcripts. After amplification of full-length *PROS1* transcripts all PCR products were ligated into sequencing vector pCR4.0. Approximately 40 clones of each cell type were analysed resulting in a TSS distribution estimate. The positional indications used throughout this article, are respective to the translational startcodon.

Target	Name[a]	Assay	Location	Primer/Probe sequence 5'-3'
PS	PS-F	QPCR	+32/+51 ex1	TGCTGGCGTGTCTCCTCCTA
	PS-R	QPCR	+109/+85 ex2	CAGTTCTTCGATGCATTCTCTTTCA
	PS-P	QPCR	+55ex1/+7ex2	CTCCCCGTCTCAGAGGCAAACTTTTTGTC
PBGD	PBGD-F	QPCR	+13/+28 ex1	GGCAATGCGGCTGCAA
	PBGD-R	QPCR	+25/+43 ex2	GGGTACCCACGCGAATCAC
	PBGD-P	QPCR	+30ex1/+23ex2	CTCATCTTTGGGCTGTTTTCTTCCGCC
PS	PSstart	GR	-1/-21	GAATTCGAAGCGCGCGGAGGCGCC
PS	PSex1	GR	+62/+33	ACGGGAAGCACTAGGAGGAGACACGCCAG
PS	PSex2	GR	+40/+21	CTTCCTAACCAGGACTTGTG

Table 1 Primer and probe sequences. F, forward primer; R, reverse primer; P, probe. All Taqman probes contain the reporter dye, 6-carboxy-4,7,2',7'-tetrachlorofluorescein, TET, and the quencher dye 6-carboxy-tetramethyl-rhodamine, TAMRA. GR= gene racer.

Results

PROS1 mRNA and PS protein are produced by HepG2, HuH7*, HeLa, HUVEC, Meg01 and human liver* --- PS circulates in normal plasma at a concentration of 0.33 μM (14). Hepatocytes are the largest contributor to the systemic concentration (15), but other cell types have also been shown to produce PS (16-18;20). ELISA analysis demonstrated that all cell lines used in this study produce PS (Figure 2). HepG2, and HuH7 cell lines were most productive whereas Meg01, HeLa, and HUVEC cells expressed smaller amounts of PS. Since mature megakaryocytes (platelets) are known to contain storage pools of coagulation factors in their α-granules, cell lysates were also analysed for PS antigen levels (36). Very low and comparable levels of PS antigen were found in lysates from all cell types including Meg01 (≤10 fmol per 10^7 cells). *PROS1* transcript levels were determined by QPCR of total RNA from each cell line. All cell lines were found to contain *PROS1* transcripts, indicating *de novo* production of PS. The relative level of *PROS1* transcript in the various cell types correlated with the PS protein level found in each cell type, i.e. HepG2 and HuH7 cells produced high amounts of PS protein and contained high relative levels of *PROS1* transcript.

PROS1 mRNA contains multiple start sites --- For TSS determination, full-length *PROS1* transcripts were isolated from polyA RNA samples from all cell lines and from a human liver sample. Figure 3a shows the PCR products obtained after selection and amplification of full-length *PROS1* mRNA products. Specificity of the amplified products was confirmed with a nested *PROS1* PCR, in which an internal GSP was used on the PCR mix (Figure 3b). The size of the resulting *PROS1* fragments decreased by the expected number of base pairs (62 bp). Sequence analysis of the cloned cDNAs revealed the presence of multiple TSSs in *PROS1* mRNA in all cell types. The start site distributions (Figure 3c) generated from these data revealed that four distinct start sites, located at -200, -147, -117 and -100 bp upstream from the translational start, can be defined for *PROS1*. Other start sites were found, but most were encountered only once. HUVEC and HuH7 cells do not use the TSS at -100 bp, whereas all other mRNA samples do contain this starting point. Also, HUVEC cells use the initiation site at -200 bp more frequently than other cell types. Long transcripts commencing at -340 bp, -347 bp, and -360 bp were found for Meg01, HeLa and HuH7, respectively.

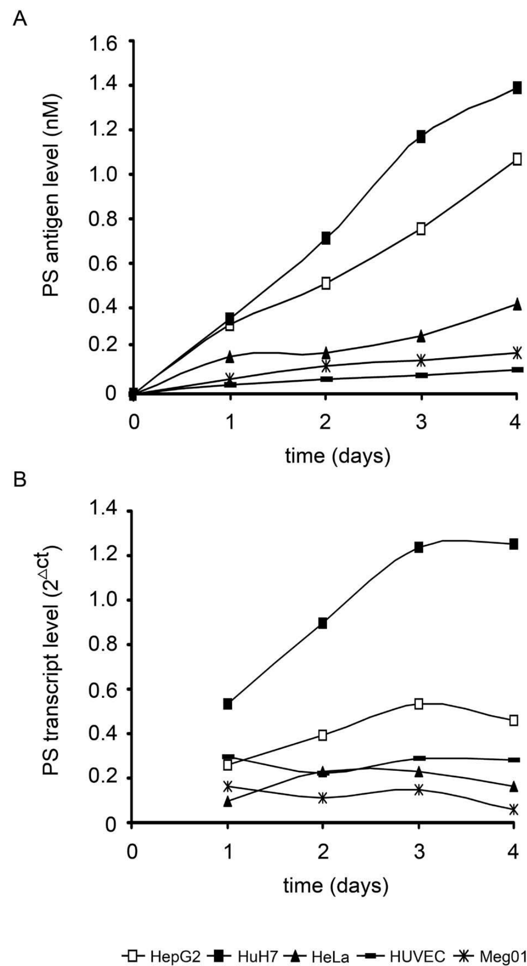
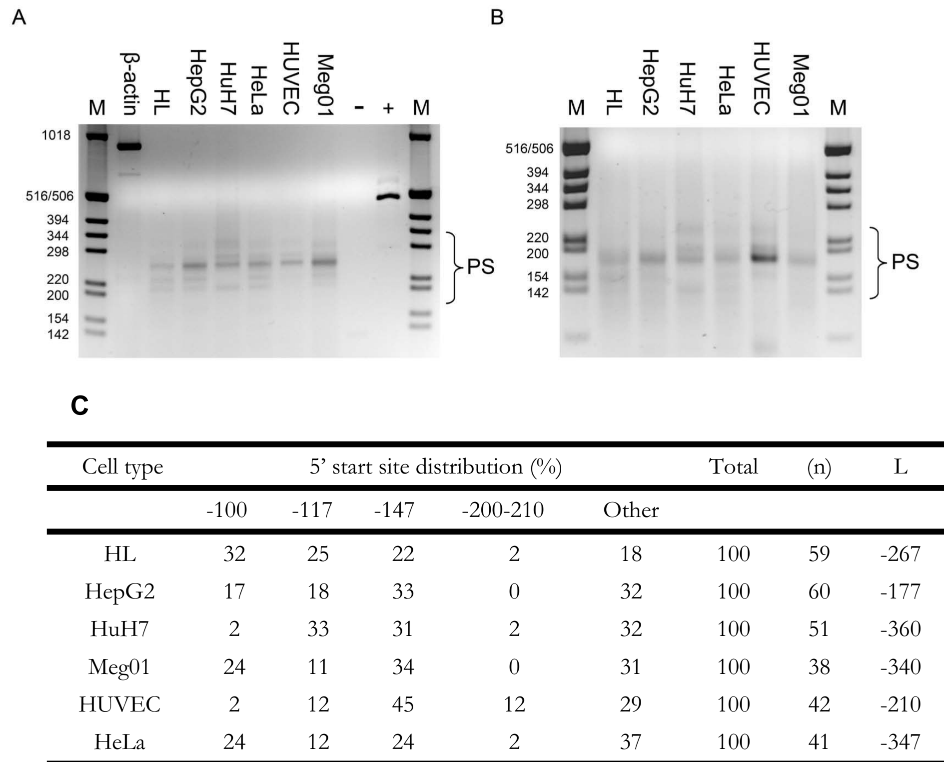


Figure 2 PS antigen and mRNA levels in various cell lines. (**A**) PS antigen levels were measured over time in media from cell culture. Cells were seeded in 12 well plates and incubated with 1 ml medium. Experiments were started at 80% confluency or, in the case of Meg01, at $1*10^6$ cells/ml. HepG2, HeLa, Meg01 and HuH7 medium included 10% FBS, HUVEC medium was serum-free. Fresh media in the appropriate dilutions were used as blanks to correct for background interference in the ELISA. Media and total RNA from triplicate samples were pooled. (**B**) PS mRNA levels are expressed as fold transcript relative to the internal standard PBGD ($2^{\Delta Ct}$).



Cell type	5' start site distribution (%)					Total	(n)	L
	-100	-117	-147	-200-210	Other			
HL	32	25	22	2	18	100	59	-267
HepG2	17	18	33	0	32	100	60	-177
HuH7	2	33	31	2	32	100	51	-360
Meg01	24	11	34	0	31	100	38	-340
HUVEC	2	12	45	12	29	100	42	-210
HeLa	24	12	24	2	37	100	41	-347

Figure 3 *PROS1* transcription is directed from multiple start sites. (**A**) Full-length *PROS1* cDNAs were amplified using the GeneRacer™ technique. M; DNA size standard, β-actin; positive 5' RACE control, -; negative PCR control, e.g. the reaction without cDNA, +; PS positive PCR control with primers spanning from exon 2 to exon 5. The PS primer is located in exon 1 at +62 to +32 (PSex1) (**B**) Nested PCR on the PCR mix which is depicted in Figure 3A. The PS primer is located in the 5' UTR at -1 to -21 (PSstart). M; DNA size standard. (**C**) Observed start site frequencies are given as percentage of the total. TSS numbering is relative to the translational startcodon. "other" indicates start sites different from the ones denoted, start sites in this category occurred no more than twice. (n) equals the number of PCR clones sequenced. L; longest transcript.

No evidence for an alternative promoter --- In 1986 Lundwall *et al* reported the presence of a *PROS1* transcript that starts in intron A (33). This could indicate the presence of an alternative promoter in intron A. The transcript has an intronic sequence of 236 bp and commences approximately 14.9 kb downstream from the first exon and 31.1 kb upstream from the second exon. This transcript is not prominently present in our HepG2 mRNA population. Figure 4 shows a 5' RACE PCR, which was conducted with a PS-specific primer in exon 2. The

products were not digested by *Mse*I, a restriction enzyme with a site in the alternative transcript that was found by Lundwall and co-workers (33). Moreover, all PCR products were digested by *Ava*II, for which restriction sites are present in exon 1.

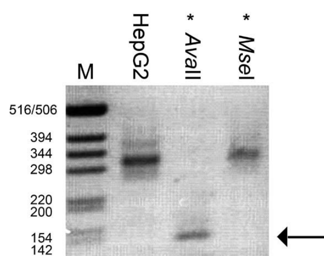


Figure 4 *PROS1* transcripts with a start in intron A are not present. Full-length *PROS1* transcripts were amplified from HepG2 mRNA using the GeneRacer™ technique. PSex2 was used as GSP for PS in the second exon. M; DNA size standard. The arrow indicates the specific digestion products after digestion with *Ava*II.

A minimal promoter of 370bp confers maximal transcriptional activity

5' deletion experiments --- Transient transfection studies were conducted in unstimulated cells with equimolar amounts of *PROS1* promoter constructs cloned upstream from the luciferase reporter gene (Figure 5a). Luciferase levels differed widely per transfected cell type and are dependent on transfection efficiency of the cell type and on the presence of the correct combination of nuclear transcription factors by which transcription is regulated. For comparison purposes, promoteractivity of the most active promoter construct in each cell type was adjusted to 100% after which the activity of the other constructs in the same cell type was normalised to this 100%. These studies pointed out that the first 370 bp of the *PROS1* 5' flanking region are sufficient for optimal promoter activity in all cell types. This activity was maintained up to a 5' region of 1062 bp. A marked difference in activity for longer constructs was observed among the various cell lines. In HuH7 and Meg01 cells promoter activity decreased only slightly with increasing construct size, whereas in HepG2 and HeLa cells promoter activity decreased more rapidly with increasing construct size.

3' deletion experiments --- Transient transfection experiments in HepG2 with *PROS1* 3' deletion reporter constructs confirmed the importance of the transcription start sites located at -100 bp and -117 bp. Deletion of the first 131 bp upstream from the startcodon decreased

promoter activity by approximately 50% (Figure 5b). The importance of the -147 bp TSS could not be confirmed in these experiments since promoter activity remained at a similar level after further deletion up to -181 bp. Transcriptional activity decreased significantly after deletion of the 5' region up to -256 bp from the startcodon, and deletion up to -806 bp completely reduced the level of activity to that of pGL3basic. Comparable results were obtained in all cell lines (results not shown). Since the -147 bp start site was one of the three most prominently found endogenous TSSs in all cell lines, we set out to validate the TSS usage in the transfected *PROS1*-luciferase constructs.

Construct start site determination --- The transcription start sites that were used in *PROS1* promoter constructs, PS370, PS1062, and PS5948 were determined after transfection in HepG2 cells. If the transfection system reflects the *in vivo* situation, the TSSs from the transfected constructs should be identical to those of the endogenous *PROS1* gene. Surprisingly, a large range of mostly unique TSSs was found for the three constructs tested (Table 2). Start sites -100 and -117, which were shown to be important sites in endogenous *PROS1* transcriptional initiation, were also used by the constructs. However, these sites were not used more frequently than the other TSSs.

Discussion

The experiments described in this study provide a more detailed understanding of the transcription start site usage during the initiation of *PROS1* transcription in various cell lines and human liver. Our results show that human liver has three prominent start sites, namely -100 bp, -117 bp and -147 bp upstream from the translational start. These TSSs were found in most cell types tested. Whereas hepatoma cell line HepG2 uses TSS -100 bp, this TSS was not found for the related cell line HuH7. Since human liver does contain this TSS we conclude that, for *PROS1* transcription, HepG2 cells are more representative of the *in situ* situation than HuH7 cells. HUVEC cells also do not use the aforementioned start site, moreover in this cell type transcription is regularly initiated from -200 bp. This TSS is rare in the other cell types suggesting that, in HUVEC, a different array of regulatory nuclear proteins may be present. Even though human liver, HepG2, HeLa, and Meg01 cell types all contain the same start sites, the contribution of the three main start sites to the total *PROS1* mRNA pool was different (Figure 3c).

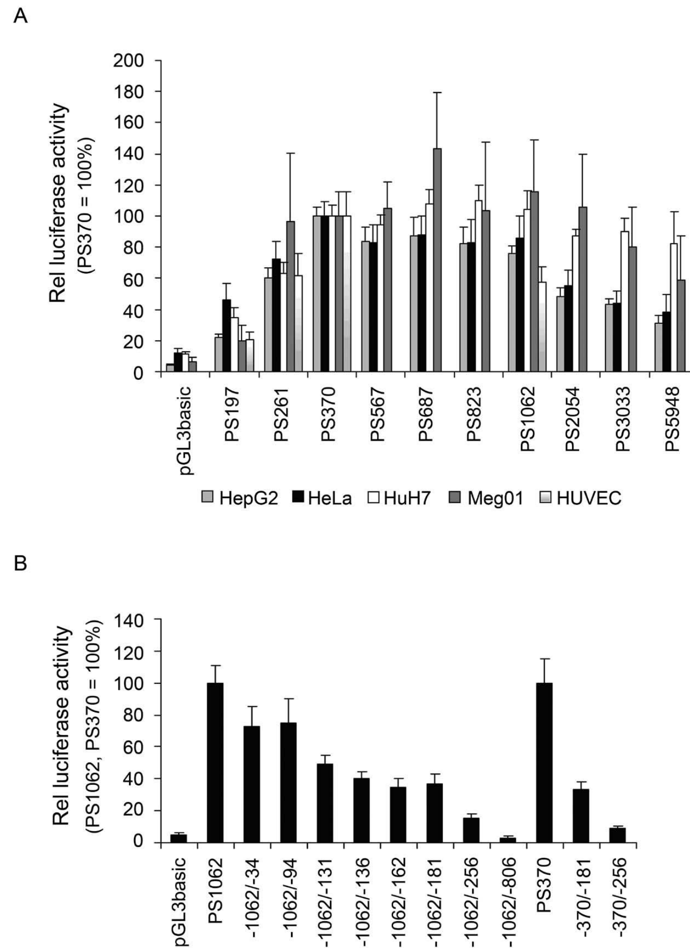


Figure 5 A minimal promoter of 370 bp is sufficient for maximal promoter activity. Cell lines were transiently transfected with the pGL3basic vector carrying *PROS1* promoter-luciferase gene constructs. Bars represent a mean of 2 triplicate assays with 2 different preparations of each plasmid (n = 12). (**A**) Luciferase activity is relative to the activity of construct PS370, which was designated as 100%. Note that the activity for PS370 was different for each cell line. Luciferase activities can therefore not be compared among cell lines, as they are also dependent on transfection efficiency in addition to promoter activity (**B**) HepG2 cells were transfected with 3' deletion constructs. Luciferase activity is relative to the activity of the 5' parent construct, i. e. PS1062 or PS370.

Construct	5' start position of cDNA				
	-1-100	-100-200	-200-300	-300-400	>-400
370	-36,-60,-73	-100,-115	-238,-242	-308,-320,-356	
1062	-59,-65,-87	-100,-119	-203,-242, -289	-308,-328,-332, -337,-395*	-412,-434, -464, -523
5948	-2	-100*,-119	-230,-238, -247, -290	-302,-318*, -332*	-424

Table 2 PS construct start site distribution. Full-length luciferase cDNAs were amplified from mRNA isolated from transfected HepG2 cells using the GeneRacer™ technique. Primer GL2 was used as GSP for luciferase mRNA. TSS numbering is relative to the luciferase translational startcodon. The start sites were all unique with exception of those with an asterisk (*), for these TSSs n=2.

Published and unpublished (NCBI database entries) *PROS1* cDNA sequences suggest the presence of multiple TSSs. To our knowledge for the first time, our results pinpoint exactly which TSSs are used for PS transcription in human liver and various cell lines. In accordance with some database entries we show that transcriptional initiation can take place from sites located more than 300 bp upstream from the translational start. However, these start sites are not used frequently. One can conclude that for GC-rich promoters, such as the *PROS1* promoter, the longest reported transcript does not necessarily represent the most abundantly used TSS.

The transcriptional regulation of the *PROS1* gene has, to date, been poorly investigated. According to a recent review this may be due to the presence of an additional exon 5' to what currently is considered exon 1 (37). Ploos van Amstel *et al* (25) identified two TSSs by primer extension analysis at -174 bp and -286 bp, respectively. Based on these findings and on the presence of a putative splice acceptor site, they postulated the presence of an alternative promoter and first exon. Also, Lundwall *et al* reported the presence of a *PROS1* transcript that starts in intron A (33). In our studies we found only uninterrupted *PROS1* sequences stretching from the 5' flanking region into exon 1. Also, restriction analysis of GeneRacer products generated with a GSP in exon 2 provided no support for the hypothesis of an alternative promoter.

At present, two DNA elements have been found to be involved in core promoter function in the transcriptional regulation from TATA-less promoters: the initiator (Inr), and the downstream promoter element (DPE)(38). The DPE is located precisely 28-30 nucleotides downstream from the transcription start site (for review see (39)). TSSs -200 bp, and -100 bp, form the center of an Inr consensus sequence (YYA$_{+1}$NWYY). However, the consensus sequence (RGWYV) for the DPE was not detected in the *PROS1* promoter at the designated positions. A schematic presentation of the TSSs and their interspecies conservation, is shown in Table 3. Phylogenetic footprinting revealed that the region surrounding the -100 bp TSS is poorly conserved, whereas the other regions have a higher degree of conservation.

Species	-200	-147	-117	-100
Human	.TATC<u>A</u>CTTCC.	.CGGC<u>T</u>GTTCC.	.GGCG<u>C</u>CGCCG.	.GCTC<u>A</u>GACCG.
Chimpanzee	.TATCACTTCC.	.CGGCTGTTCC.	.GGCGCCGCCG.	.GCTCAGACCG.
Rhesus monkey	.TATCACTTCC.	.CGGCTGTTCC.	.GGCGCCGCCG.	.GCTCAGACCG.
Dog	.TATCACTTCC.	.CGGC(*)TGTTGG.	.GGCGCCGCCG.	.CCTCGGAGAG.
Pig	.TATCACTTCC.	.CGGCTCGGCA.	.GGCGCCGCTG.	.CTTCAGAGAG.
Mouse	.CACCGCTTCC.	.TGGCTGCTCC.	.GG.GCCGCCG.	.GCTGGGAGAG.
Rat	.CGCCGCTTC..	.TGACTGCTCC.	.GG.GCCGCAG.	.GTTGGGAGAG.
Cow	.TATCACTTCC.	.CGGCTCCGCT.	.GGCGCCCCCG.	.CCTTAC.TAG.

Table 3 Phylogenetic footprint for the most prominent *PROS1* TSSs. Gaps introduced to maximize alignment are indicated by dots. The TSS is underlined.

A region up to 370 bp upstream from the translational start is sufficient for optimal promoter expression in our model transfection system. We show that this activity is maintained up to a length of 1062 bp upstream. The subtle differences between the cell lines in these transfection studies do not allow a firm conclusion regarding tissue-specific regulation of the *PROS1* promoter under basal culture conditions. However, it can be concluded that the promoter region apparently does not contain additional upstream regulatory elements, e.g. enhancers and repressors that can be detected in our model system. Surprisingly, only a 60% decrease of promoter activity was shown after deletion of the three most prominent start sites in the 3' deletion experiments. This led us to investigate the TSS usage in three of the *PROS1* promoter constructs that were used in our transfection studies. The TSSs at -100 bp and -117

bp were observed for all constructs but the -147 bp and -200 bp start sites were not found amongst the cloned mRNAs derived from the transfected constructs. Moreover, even though similar TSSs were found for the three constructs that were tested, most identified transcripts were rare, indicating a relaxation of the transcriptional machinery (Table 2). A reduction to 10% of original promoter activity was obtained after deletion of the first 256 bp proximal to the translational start. The deleted region contained all of the endogenously important TSSs (-100, -117, -147 and -200). However, the absence of these sites is most likely not the sole cause of the abrogation of promoter activity. By deletion of this region also essential transcription factor binding sites may have been deleted. For instance, Tatewaki and co-workers (40) have recently shown transcription factor Sp1 to bind within the deleted region, at -247 to -238 bp. Since Sp1 is a major component in the transcriptional regulation from many genes with GC-rich promoters (41), deletion of this site may well be, at least in part, responsible for abrogation of promoter activity. A similar argument probably applies to the PS-1062/-806 construct, which has lost all promoter activity.

Our findings suggest that the mechanism of transcriptional initiation in this model system differs from the *in vivo* situation. A possible explanation for this phenomenon may be that transfected DNA is in plasmid form and not in a chromosomal form. It therefore may lack necessary structural elements that are present in normal genomic DNA. Based on our results, researchers should express great caution to use this model system for determining TSSs. We believe that even though transcriptional initiation is directed differently, i.e. more relaxed, this system remains a valuable tool for examining promoter regulation as the binding of essential transcription factors and their role in transcriptional activation is in general not affected as has been shown in numerous publications. However, it is clear from our results that transient transfection of luciferase reporter gene constructs may not be a suitable tool for dissecting subtle differences in promoter activity and TSS usage.

Acknowledgements

We would like to thank Jos Grimbergen and Erna Peters from the TNO Center for Prevention and Health for their generous donation of HUVEC cells and for technical assistance with HUVEC transfection experiments. This work was financially supported by the Dutch Thrombosis Foundation (grant application number TSN 98.002).

References

1. Comp PC, Nixon RR, Cooper MR, Esmon CT. Familial protein S deficiency is associated with recurrent thrombosis. J.Clin.Invest 1984;74:2082-8.
2. Comp PC, Esmon CT. Recurrent venous thromboembolism in patients with a partial deficiency of protein S. N.Engl.J.Med. 1984;311:1525-8.
3. Schwarz HP, Fischer M, Hopmeier P, Batard MA, Griffin JH. Plasma protein S deficiency in familial thrombotic disease. Blood 1984;64:1297-300.
4. Kamiya T, Sugihara T, Ogata K, Saito H, Suzuki K, Nishioka J *et al.* Inherited deficiency of protein S in a Japanese family with recurrent venous thrombosis: a study of three generations. Blood 1986;67:406-10.
5. Engesser L, Broekmans AW, Briët E, Brommer EJ, Bertina RM. Hereditary protein S deficiency: clinical manifestations. Ann.Intern.Med. 1987;106:677-82.
6. Bosson JL, Francois P, Pernod G, Carpentier P. [Venous thromboembolic disease in children related to transient protein S deficiency following varicella]. Presse Med. 1995;24:415.
7. De Stefano V, Finazzi G, Mannucci PM. Inherited thrombophilia: pathogenesis, clinical syndromes, and management. Blood 1996;87:3531-44.
8. Walker FJ. Regulation of activated protein C by a new protein. A possible function for bovine protein S. J.Biol.Chem. 1980;255:5521-4.
9. Fulcher CA, Gardiner JE, Griffin JH, Zimmerman TS. Proteolytic inactivation of human factor VIII procoagulant protein by activated human protein C and its analogy with factor V. Blood 1984;63:486-9.
10. Walker FJ. Protein S and the regulation of activated protein C. Semin.Thromb.Hemost. 1984;10:131-8.
11. Walker FJ, Chavin SI, Fay PJ. Inactivation of factor VIII by activated protein C and protein S. Arch.Biochem.Biophys. 1987;252:322-8.
12. Dahlbäck B, Stenflo J. High molecular weight complex in human plasma between vitamin K-dependent protein S and complement component C4b-binding protein. Proc.Natl.Acad.Sci.U.S.A 1981;78:2512-6.
13. Dahlbäck B. Inhibition of protein Ca cofactor function of human and bovine protein S by C4b-binding protein. J.Biol.Chem. 1986;261:12022-7.
14. Griffin JH, Gruber A, Fernandez JA. Reevaluation of total, free, and bound protein S and C4b-binding protein levels in plasma anticoagulated with citrate or hirudin. Blood 1992;79:3203-11.
15. Fair DS, Marlar RA. Biosynthesis and secretion of factor VII, protein C, protein S, and the Protein C inhibitor from a human hepatoma cell line. Blood 1986;67:64-70.
16. Schwarz HP, Heeb MJ, Wencel-Drake JD, Griffin JH. Identification and quantitation of protein S in human platelets. Blood 1985;66:1452-5.
17. Ogura M, Tanabe N, Nishioka J, Suzuki K, Saito H. Biosynthesis and secretion of functional protein S by a human megakaryoblastic cell line (MEG-01). Blood 1987;70:301-6.
18. Fair DS, Marlar RA, Levin EG. Human endothelial cells synthesize protein S. Blood 1986;67:1168-71.
19. Stern D, Brett J, Harris K, Nawroth P. Participation of endothelial cells in the protein C-protein S anticoagulant pathway: the synthesis and release of protein S. J.Cell Biol. 1986;102:1971-8.

20. Malm J, He XH, Bjartell A, Shen L, Abrahamsson PA, Dahlbäck B. Vitamin K-dependent protein S in Leydig cells of human testis. Biochem.J. 1994;302(Pt 3):845-50.
21. Maillard C, Berruyer M, Serre CM, Dechavanne M, Delmas PD. Protein S, a vitamin K-dependent protein, is a bone matrix component synthesized and secreted by osteoblasts. Endocrinology 1992;130:1599-604.
22. Stitt TN, Conn G, Gore M, Lai C, Bruno J, Radziejewski C *et al.* The anticoagulation factor protein S and its relative, GAS6, are ligands for the Tyro 3/Axl family of receptor tyrosine kinases. Cell 1995;80:661-70.
23. Shimokawa T, Yamamoto K, Yamafuji E, Kojima T, Saito H. Expression of protein S in the murine heart and cultured mouse cardiomyocytes is down-regulated by cytokines. Thromb.Haemost. 2001;86:623-9.
24. Schmidel DK, Tatro AV, Phelps LG, Tomczak JA, Long GL. Organization of the human protein S genes. Biochemistry 1990;29:7845-52.
25. Ploos van Amstel HK, Reitsma PH, van der Logt CP, Bertina RM. Intron-exon organization of the active human protein S gene PS alpha and its pseudogene PS beta: duplication and silencing during primate evolution. Biochemistry 1990;29:7853-61.
26. Edenbrandt CM, Lundwall A, Wydro R, Stenflo J. Molecular analysis of the gene for vitamin K dependent protein S and its pseudogene. Cloning and partial gene organization. Biochemistry 1990;29:7861-8.
27. Gandrille S, Borgel D, Ireland H, Lane DA, Simmonds R, Reitsma PH *et al.* Protein S deficiency: a database of mutations. For the Plasma Coagulation Inhibitors Subcommittee of the Scientific and Standardization Committee of the International Society on Thrombosis and Haemostasis. Thromb.Haemost. 1997;77:1201-14.
28. Gandrille S, Borgel D, Sala N, Espinosa-Parrilla Y, Simmonds R, Rezende S *et al.* Protein S deficiency: a database of mutations--summary of the first update. Thromb.Haemost. 2000;84:918.
29. Pollak ES, Hung HL, Godin W, Overton GC, High KA. Functional characterization of the human factor VII 5'-flanking region. J.Biol.Chem. 1996;271:1738-47.
30. Huang MN, Hung HL, Stanfield-Oakley SA, High KA. Characterization of the human blood coagulation factor X promoter. J.Biol.Chem. 1992;267:15440-6.
31. Ploos van Amstel HK, van der Zanden AL, Reitsma PH, Bertina RM. Human protein S cDNA encodes Phe-16 and Tyr 222 in consensus sequences for the post-translational processing. FEBS Lett. 1987;222:186-90.
32. Hoskins J, Norman DK, Beckmann RJ, Long GL. Cloning and characterization of human liver cDNA encoding a protein S precursor. Proc.Natl.Acad.Sci.U.S.A 1987;84:349-53.
33. Lundwall A, Dackowski W, Cohen E, Shaffer M, Mahr A, Dahlbäck B *et al.* Isolation and sequence of the cDNA for human protein S, a regulator of blood coagulation. Proc.Natl.Acad.Sci.U.S.A 1986;83:6716-20.
34. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25:3389-402.
35. Deutz-Terlouw PP, Ballering L, van Wijngaarden A, Bertina RM. Two ELISA's for measurement of protein S, and their use in the laboratory diagnosis of protein S deficiency. Clin Chim Acta 1990;186:321-4.
36. Harrison P, Cramer EM. Platelet alpha-granules. Blood Rev. 1993;7:52-62.

37. Rezende SM, Simmonds RE, Lane DA. Coagulation, inflammation, and apoptosis: different roles for protein S and the protein S-C4b binding protein complex. Blood 2004;103:1192-201.
38. Smale ST, Baltimore D. The "initiator" as a transcription control element. Cell 1989;57:103-13.
39. Smale ST, Kadonaga JT. The RNA polymerase II core promoter. Annu.Rev.Biochem. 2003;72:449-79.
40. Tatewaki H, Tsuda H, Kanaji T, Yokoyama K, Hamasaki N. Characterization of the human protein S gene promoter: a possible role of transcription factors Sp1 and HNF3 in liver. Thromb.Haemost. 2003;90:1029-39.
41. Suske G. The Sp-family of transcription factors. Gene 1999;238:291-300.