YY1 binding within the human HSD3B2 gene intron 1 is required for maximal basal promoter activity: identification of YY1 as the 3β1-A factor

D M Foti and J K Reichardt1

Department of Biochemistry and Molecular Biology, Institute for Genetic Medicine, USC Keck School of Medicine, Los Angeles, California 90033, USA

1Department of Preventive Medicine, Institute for Genetic Medicine, USC Keck School of Medicine, Los Angeles, California 90033, USA

(Requests for offprints should be addressed to J K Reichardt; Email: reichard@usc.edu)

Abstract

The oxidation and isomerization of 3β-hydroxy-5-ene steroids into keto-4-ene steroids, a pivotal step in the synthesis of all hormonal steroids, is catalyzed by several isoforms of 3β-hydroxysteroid dehydrogenase. In humans, two highly homologous isoforms exist, type I expressed by the HSD3B1 gene in peripheral tissues, and type II expressed by the HSD3B2 gene in steroidogenic organs. Previously, it was shown that the HSD3B1 gene 3β1-A element, encompassing 24 nucleotides of intron 1 not perfectly conserved between the two genes and overlapping with a conserved TG box, contributes to maximal basal promoter activity by binding the ubiquitous and unidentified 3β1-A transcription factor. In this study for the first time we report that similarly, the HSD3B2 gene intron 1 is required for maximal basal promoter activity in reporter gene analyses, as lack of intron 1 results in a 4- to 10-fold reduction in promoter activity. Mutational analysis in gel shift assays revealed that the 3β1-A factor binds both the HSD3B2 and HSD3B1 gene intron 1 by requiring only seven nucleotides of a conserved segment within the 3β1-A element. By competition analysis and use of anti-YY1 antibody in both gel shift and Western blot experiments, we identified the 3β1-A protein as the ubiquitous transcription factor YY1. In addition, we have characterized another similar YY1 binding site differently located with respect to the 3β1-A element in both genes. Deletion and mutational analysis in transient transfections experiments revealed that contrarily to as previously shown for the HSD3B1 gene, lack of YY1 binding to the type II 3β1-A element only results in a marginal reduction of basal promoter activity. Instead, YY1 binding to the second site, placed 35 bp downstream from the 3β1-A element, strongly activates the HSD3B2 gene basal promoter activity, as preventing YY1 binding to this region caused a 50% decrease of basal transcription. Complete abrogation of YY1 binding within type II intron 1 resulted in a gene reporter activity identical to a reporter construct lacking the whole intron 1. These results designate YY1 as the factor responsible for the intron 1-mediated boost of the HSD3B2 gene basal promoter activity. Similarities and dissimilarities between YY1 binding within the HSD3B1 and HSD3B2 gene intron 1 are discussed involving the conserved intron 1 TG box, that suggests different mechanisms are implicated in the YY1-mediated stimulation of these two genes basal promoter activity.

Journal of Molecular Endocrinology (2004) 33, 99–119

Introduction

The enzyme 3β-hydroxysteroid dehydrogenase (3β-HSD) catalyzes a fundamental step in the synthesis of all hormonal steroids in many tissues and organs (Labrie et al. 1992), which play a crucial role in the differentiation, development, growth, and physiological function of most human tissues. It is also one of the two enzymes that start the degradation of dihydrotestosterone (DHT), the most powerful androgen (Bartsch et al. 1990), thus affecting the availability of this hormone in the regulation of androgen-responsive genes. Two isoforms are known in humans, type I and type II 3β-HSD, encoded by the HSD3B1 and HSD3B2 gene respectively (Luu-The et al. 1989, Lachance et al. 1990, 1991, Rheaume et al. 1991). The two genes are closely linked on the short arm of
The recent discovery of a binding site for steroidogenic factor 1 (SF-1) at nucleotide −64 to −56 of the HSD3B2 gene promoter (Leers-Sucheta et al. 1997), which is not conserved in the HSD3B1 gene, might explain the restricted expression of the HSD3B2 gene to steroidogenic organs. SF-1 had indeed earlier been identified as a tissue-specific transcription factor that regulates all the steroidogenic P-450 genes in the adrenal cortex and gonads (Lala et al. 1992, Morohashi et al. 1992). Interestingly, a study by Guerin et al. (1995) identified the 3β1-A element, a positive regulatory cis-acting element located within the HSD3B1 gene intron 1 (Fig. 1) that is required for reaching maximal promoter activity in transient transfection assays. This region, encompassing 24 nucleotides protected in DNase footprinting assay, was shown to strongly bind an unidentified ubiquitously expressed nuclear protein named 3β1-A factor, whose molecular mass was estimated as 37 kDa by UV-crosslinking experiments. Mutational analysis revealed that the four Gs that lay at the 3′ end of the protected region, shared with an overlapping TG box (Fig. 1), are essential for binding of the 3β1-A factor to the regulatory element, and their mutation resulted in a 3- to 6-fold lower expression of a reporter gene. As Sp1 was found to bind the TG box very weakly, the 3β1-A protein was attributed a transactivator role. Stressing the importance of the 3β1-A element in enhancing basal transcription, Guerin and collaborators also showed that a double-stranded oligomer bearing the DNA sequence of the HSD3B1 gene 3β1-A element cloned in front of the mouse p12 gene promoter was able to stimulate transcription of the CAT reporter gene from 2- to 5-fold in human choriocarcinoma JEG-3 and human adrenal cortex adenocarcinoma SW-13 cells respectively.

Because of the importance of the type II 3β-HSD activity in steroidogenesis, we were prompted to investigate whether or not the HSD3B2 gene intron 1 has any functional relevance, as in the case of the HSD3B1 gene. We here report that intron 1 is required by the upstream promoter of the HSD3B2 gene to achieve the maximal basal level of transcription, as deletion analysis in transient transfection assays revealed that removal of intron 1 results in a drastic reduction of luciferase reporter activity. Despite mismatches in the region, we have ascertained the presence of a 3β1-A element in the HSD3B2 gene intron 1. Furthermore, we have
identified the 3β1-A protein to be the multifactorial transcription factor YY1 (Shi et al. 1997), and demonstrate that, in addition to the 3β1-A element, YY1 also recognizes a second binding site within intron 1, whose position is not conserved between the two genes. Through mutational analysis and co-expression of YY1 in transfection experiments we were able to correlate YY1 with the boost in basal activity mediated by type II intron 1, and identified the second YY1 binding site, not the 3β1-A element, as a major target for YY1-mediated boost of HSD3B2 gene basal promoter activity.

In addition, we provide evidence that Sp3 binds the TG box in the intron 1 of both genes and that the binding efficiency is much higher for the type I TG box. We found the second YY1 binding site in the binding efficiency is much higher for the type I TG box in the intron 1 of both genes and that boost of HSD3B2 gene basal promoter activity.

We demonstrated that, in addition to the 3β1-A element, YY1 also recognizes a second binding site within intron 1. Although we have not investigated the functional relevance of those findings for the HSD3B1 gene, our results point to a role for YY1 in the basal expression of 3β-HSD activity, and suggest that different mechanisms may be involved in the maintenance of the HSD3B1 and HSD3B2 gene basal activity under experimental conditions, as discussed.

Materials and methods

Plasmid constructions

The human HSD3B2 gene 5′-flanking and 5′-UTR (from −1265 to +193) was amplified by PCR from human genomic DNA isolated from blood using the forward primer 5′-CCGACGCCGTTAATAAAA CATTTAAGCCATAATAAAA-3′ and the reverse primer 5′-GGGCTCGAGACCCAGAAGA GGGCTAAAAAC-3′. The amplification product was digested with MluI and XhoI and cloned in pGL3-Basic vector multiple cloning site. The insert nucleotide sequence was verified by automated DNA sequencing using the forward RV3 primer 5′-CTAGAAAAATAAGGCTGT CCC-3′, and reverse GL2 primer 5′-CATTATATGTT TTTTGCCGTCTTTCC-3′ from Promega, which respectively match upstream and downstream to the pGL3-Basic vector multiple cloning site.

Site-directed mutagenesis

Mutations were introduced into the reporter plasmids using the QuikChange Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA, USA) following the manufacturer’s instructions and using the same mutant oligomers that are represented in the figures that accompany this paper. Mutations were confirmed by DNA automated sequencing (ABI PRISM 3100 Genetic Analyzer; Applied Biosystems, Foster City, CA, USA). Each sequence-confirmed mutant insert was released with the appropriate cloning restriction endonucleases, and inserted into similarly digested pGL3-Basic, to ensure the plasmid used in transfections did not carry any PCR-induced nucleotide substitutions within the vector backbone.

Gel shift assays

Crude nuclear extracts were prepared as previously described by Dent & Latchman (1993). Oligonucleotides for radiolabeled probes and competitor DNAs were obtained from Invitrogen Corporation and Integrated DNA Technologies (Coralville, IA, USA). Appropriate sense and antisense strands were annealed in Tris 10 mM pH 8·0 by heating at 85 °C for 3 min and slowly cooling down to 25 °C. Double-stranded oligomeric probes were 5′ end labeled with [γ-32P]ATP (5000 Ci/mmol) using T4 polynucleotide kinase (New England Biolabs, Beverly, MA, USA) according to the manufacturer’s instructions. Labeled oligomers were separated from unreacted nucleotides by centrifugation through ProbeQuant G-50 Micro Columns (Amersham Biosciences). One, 3 or 5 μg nuclear proteins were pre-incubated with 1 μg poly (dI-dC)-poly(dI-dC) (Sigma-Aldrich) in 20 μl binding buffer (Hepes 20 mM pH 7·9, KCl 50 mM, MgCl2 5 mM, dithiothreitol 0·5 mM, and 4% glycerol) for 5 min at room temperature. The

www.endocrinology.org
labeled probes were then added (100 000 c.p.m., approximately 0·1–0·2 ng) and complexes were allowed to form at room temperature for 15 min. The binding reactions were loaded onto a 4·5% non-denaturing polyacrylamide gels containing 2·5% glycerol and subjected to electrophoresis in 0·5 × TBE buffer at 17–20 mA for about 3 h at room temperature. The gel was dried and either autoradiographed overnight with an intensified screen or exposed to a Storage Phosphor Screen (Amersham Biosciences) and analyzed with a STORM 840 densitometer (Molecular Dynamics, Sunnyvale, CA, USA). The specificity of protein–DNA complexes was analyzed by including 100-fold molar excess of homologous or non-homologous unlabeled competitor DNA into the binding reactions during the pre-incubation step prior to addition of the probe. To identify nuclear factors that bind to the probes, 1 µl specific antibody (rabbit polyclonal anti-YY1, sc-1730; rabbit polyclonal anti-Sp3, sc-644 X; rabbit monoclonal anti-Sp1, sc-420 X; anti-Sp2, sc-643 X; Santa Cruz Biotechnology, Santa Cruz, CA, USA) was incubated with the binding mix for 1 h at 4 °C before addition of the labeled oligomers.

**In vitro transcription/translation assay**

*In vitro*-translated YY1 was obtained with the TNT T7/T3 Coupled Reticulocyte Lysate System from Promega. To obtain the template DNA plasmid, YY1 cDNA was released from pCMV-YY1 (gift of Dr Shi, Harvard University) by digesting with BamH1 and Kpn1 restriction endonucleases and cloned into pCMX-L1 vector similarly digested. The *in vitro* coupled transcription/translation reaction was carried out following the manufacturer’s instructions in the presence of T3 RNA Polymerase and 1 µg non-linearized template. The negative controls were obtained performing a similar reaction in the absence of DNA template, or in presence of 1 µg pCMX-L1 backbone vector.

**Acrylamide capture of 3β2 oligomer-bound complexes**

Oligo A-3β2, the 3β2 upper strand oligonucleotide modified by addition of an Acrydite moiety (Mosaic Technologies, Waltham, MA, USA) at its 5’ prime end was obtained from Integrated DNA Technologies. The procedure was carried out as described to Nelson *et al.* (2002) with few modifications: 1 µM A-3β2 oligo was annealed with the same amount of the unmodified lower strand and incubated with 150 µg SW-13 cell nuclear extract or 0·5 mg LNCaP nuclear extract under the same binding conditions herein described for the electrophoretic mobility shift assay (EMSA), scaling up the reaction volume accordingly to the amount of proteins and oligomers. After 15 min of incubation, the duplex molecules were immobilized to acrylamide as previously described (Kennet *et al.* 1997) in the 1 cm large wells of a non-denaturing 4·5% acrylamide gel (20 × 20 cm), and the unbound proteins were separated from the bound complexes by electrophoresis under the same conditions used for the EMSAs. The polyacrylamide fragments containing the covalently linked A-3β2/protein complexes were excised, and the proteins recovered by triturating the band in extraction buffer (Tris 50 mM pH 7·5, SDS 0·1%, EDTA 0·25 mM and 2·5% glycerol). Samples were kept under rotation overnight at 4 °C. The slurry was filtered through a 0·22 µm Ultrafree-MC Centrifugal Filter Unit and concentrated by a Microcon Centrifugal Filter Device (Millipore Corporation, Bedford, MA, USA). Typically, about 0·8 ml extraction buffer was reduced to 50 µl.

**Western blot analysis**

Protein concentrations were determined by the BioRad protein assay. Ten micrograms of nuclear proteins from SW-13 and LNCaP cells, and the same amount of isolated A-3β2-bound proteins, along with 10 µl *in vitro* YY1 were diluted with 2 × SDS loading buffer and loaded onto a 12% pre-cast GeneMate Endurance acrylamide gels (ISC BioExpress, Kaysville, UT, USA). Proteins were transferred to Hybond-P PVDF transfer membrane (Amersham Pharmacia Biotech) by a semi-dry procedure. Blots were probed with the same polyclonal antiserum against YY1 (1:500) or against Sp3 (1:5000) used in the gel shift experiments as primary antibodies, and horseradish peroxidase-conjugated goat anti-rabbit antibodies (sc-2030; Santa Cruz) were used as secondary antibodies. Bands were visualized using the Western blotting Chemiluminescence Luminol Reagent (Santa Cruz Biotechnology, Santa Cruz, CA, USA). Molecular mass standards recognized by the secondary antibody were used (sc-2031; Santa Cruz Biotechnology).
Cell cultures and transfections

All medium and sera were provided by Gibco BRL (Life Technology). SW-13 cells were grown in IMEM medium, and LNCaP cells in RPMI medium, both supplemented with 10% commercial fetal bovine serum (FBS), 2 mM glutamine, and 100 IU/ml penicillin and 100 µg/ml streptomycin. COS-7 cells were grown in DMEM medium supplemented with 5% FBS. Cells were maintained at 37 °C with 5% CO₂. To compare the luciferase reporter activity of plasmids pGL3-HSD2(–517/+193) and pGL3-HSD2(–517/+23), SW-13 and COS-7 cells were seeded in six-well plates at a density of 6 × 10⁵ cells/well and 1·6 × 10⁵ cells/well respectively. Sixteen hours later cells were transfected in triplicate wells by adding in each well 3·5 µg reporter plasmid, 0·5 µg pCMXβGal and 1 µg empty pCMX-L1 vector in presence of Lipofect-Amine 2000 (Invitrogen, Life Science) accordingly to the manufacturer’s instructions (a ratio DNA/Lipofect-Amine 1:1·5 was used). Cells were harvested 24 h later, and cell extracts were assayed for the luciferase activity. To test the mutant constructs reported in Fig. 7, SW-13 cells were seeded in 12-well plates at a density of 3 × 10⁵ cells/well and transfected in duplicates with 2 µg DNA and 3 µl Lipofect-Amine (1:1·5) per well. When tested alone, reporter plasmids were transfected at the concentration of 1·8 µg/well bringing the DNA amount up to 2 µg with 0·2 µg pCMXβGal to monitor transfection efficiency. Co-transfection experiments with YY1 expression vector were carried out with 1 µg/well reporter and 0·5 or 1 µg/well of expression plasmid, keeping the DNA amount constant by adding pCMX-L1 empty vector. Because of the transcription factor interference with beta-galactosidase expression vs the control samples consisting of reporter plasmid alone, no pCMXβGal was added in co-transfection experiments, and the luciferase activity was instead normalized for the protein content, when necessary. Luciferase activity was measured for 20 s in a luminometer using the Luciferase Assay System (Promega).

Results

The HSD3B2 gene intron 1 is required for sustaining the basal promoter activity

It was previously shown that deletion of intron 1, but not exon 1, reduced the HSD3B1 gene basal promoter activity from 3- to 6-fold (Guerin et al. 1995). To test whether intron 1 sequences play a similar role in the HSD3B2 gene basal promoter activity, we generated plasmid pGL3-HSD2(–517/+193), which expresses the luciferase reporter gene under the control of the HSD3B2 gene sequence from nucleotide –517 to nucleotide +193, thus encompassing exon 1 and intron 1, and plasmid pGL3-HSD2(–517/+23), which lacks intron 1 (Fig. 1B). As also intended for a study aiming at characterizing polymorphisms that occur within intron 1, which might affect the splicing of the intron, our promoter constructs encompassing the first intron include the conserved splicing portion of the 3’ splice junction with exon 2 up to 12 nucleotides (+182/+193). By RT-PCR we confirmed that including the first 12 nucleotides of exon 2 results in intron 1 being properly spliced out in COS-7 cells (result not shown). When pGL3-HSD2(–517/+193) and pGL3-HSD2(–517/+23) were compared in transient transfection experiments with pGL3 empty vector as a negative control in SW-13 cells (Fig. 1B), the promoter activity of the construct bearing intron 1 was 23-fold higher than the negative control. The activity of the construct deleted of intron 1 was stimulated only 4·5-fold over the negative control, with a loss of promoter activity as much as 4·7-fold with respect to the promoter activity of the full-length construct. In COS-7, pGL3-HSD2(–517/+193) promoter activity was stimulated 8-fold vs control, while the basal promoter activity of the intron-less construct was only stimulated 1-fold (Fig. 1B). In the steroidogenic SW-13 cells, derived from a small-cell carcinoma of the adrenal cortex (Leibovitz et al. 1973), the gene reporter activity was 3- to 4-fold higher than in COS-7 cells, a non-steroidogenic cell line, which is likely due to tissue specificity (Fig. 1B). However, the fact that deletion of intron 1 resulted in a drastic reduction in basal transcription in both steroidogenic and non-steroidogenic cell lines showed that similarly to the HSD3B1 gene the mechanism involved is not strictly tissue-specific.

The HSD3B2 intron 1 contains a 3’1-A element

To determine if a region with functional similarities to the HSD3B1 gene 3’1-A element might be responsible for the stimulating effect intron 1 demonstrated in transfection experiments, we
Figure 1 The HSD3B2 gene intron 1 is required for maximal promoter basal activity. (A) A schematic representation of the HSD3B2 gene organization and partial sequence alignment of the HSD3B1 and HSD3B2 intron 1 is shown. Open boxes indicate untranslated exon sequences as well as the 5’ and 3’ untranslated regions; solid boxes indicate the coding exons. Roman numbers identify each exon. The transcription initiation start in exon 1 is shown. In the sequence alignment, the nucleotides part of the control elements previously identified in the HSD3B1 gene intron 1 by Guerin et al. (1995) are wrapped by a shaded box and the name of each element is shown above each box. The nucleotides previously identified as protected in DNAse footprinting experiments are underlined. Small letters evidence nucleotide differences between the two aligned sequences. Dotted lines replace missing nucleotides. (B) SW-13 cells (6x10^5 cells/well) and Cos-7 cells (1x10^5 cells/well) were seeded in six-well plates and transfected with pGL3-HSD2(−517/+193) and pGL3-HSD2(−517/+23) as described in Materials and Methods. Cells were harvested after 24 h. The promoter activity is reported as relative light units (rlu, luciferase/β-galactosidase) as the mean of two independent experiments each performed with duplicate samples, and is expressed as the percentage of control vector pGL3 in SW-13 cells. The standard deviations are shown. A schematic representation of each reporter construct is shown next to the corresponding activity bar. Negative numbers indicate the first nucleotide of the HSD3B2 gene 5’ flanking region included in the construct, positive numbers identify the nucleotidic delimitation of the untranslated regions included in each construct. Open boxes schematize untranslated coding regions. The transcription initiation start (+1) is shown. A straight bold line represent the intron 1 sequence, two interconnected lines indicate deletion of intron 1.
designed oligo 3β1 (Fig. 2C), which matches the HSD3B2 gene intron 1 region corresponding to the 3β1-A element of the HSD3B1 gene. When tested with nuclear extracts from SW-13, HeLa, COS-7 and LNCaP cells in a gel shift assay, 32P-labeled 3β2 oligo formed a complex of identical electrophoretic mobility that was competed by 25-fold molar excess of cold 3β2 oligo (Fig. 2A, left panel). When 3β2 was compared with 3β1, an oligo matching the 3β1-A element (Fig. 2C) as previously explored by Guerin et al. (1995), each probe was shifted into a single major complex of identical intensity and electrophoretic mobility upon incubation with cell nuclear extracts (shown with SW-13 cells in Fig. 2A, middle panel, lanes 1 and 5). The complex was competed by 100-fold molar excess of the same unlabeled oligo (Fig. 2A, middle panel, lanes 2 and 6), but not competed by 100-fold molar excess of an oligo containing the Sp1 consensus (Fig. 2A, middle panel, lanes 3 and 7), indicating that the protein forming the complex specifically recognizes the probe. Guerin et al. (1995) had previously established that the appearance of the ubiquitous 3β1-A band requires the integrity of the stretch of four Gs shared by the 3β1-A element and TG box, as well as the presence of zinc ions in the binding reactions. To find out if the 3β1-A factor was the protein responsible for the identical mobility complex formed by the 3β1 and 3β2 probes, we tested two mutant oligos in which the four Gs were replaced with As (3β1m and 3β2m, Fig. 2C) as previously explored by Guerin et al. The two mutant oligos were unable to form the complex when used as a probe (Fig. 2A, middle panel, lanes 4 and 8), which indicates that both the complex formed by 3β1 and the complex formed by 3β2 require the G4 stretch shared with the overlapping TG box, as expected for the 3β1-A protein. When the 3β1 probe was competed by 100-fold molar excess of cold Sp1, 3β2 and 3β2m oligos, only the competition with 3β2 oligo prevented formation of the complex, as seen in Fig. 2A with LNCaP cell nuclear extract to show ubiquity of the 3β1 complex (Fig. 2A, right panel, compare lane 4 with lanes 1, 2 and 3). This result indicates that the protein forming the specific complex with 3β1 also binds to 3β2, and the binding is dependent on the integrity of the G4 stretch. To further investigate the authenticity of the complex as the 3β1-A protein, binding reactions were also performed in presence of EDTA 10 mM with and without 5 mM ZnCl2. The complex was not formed in presence of EDTA, an ion chelator (Fig. 2A, right panel, lane 5), and it was restored when Zn2+ was added to overcome the EDTA (Fig. 2A, right panel, lane 6), suggesting that the protein forming the complex requires zinc ions to bind the 3β1 probe as expected from the 3β1-A protein. Identical results were obtained when cross-competition of 3β1 and 3β2 oligonucleotides and Zn2+-dependence were tested with SW-13, COS-7, HeLa, LNCaP and PC-3 cells (not shown). Considered the identical electrophoretic mobility of the single complex formed specifically by 3β1 and 3β2, considered that both complexes are sensitive to the same changes in the G4 stretch as the 3β1-A protein, and that 3β2 oligo was able to compete for the Zn2+-dependent band formed by 3β1, it is likely that both probes are bound by the same Zn2+-dependent protein. As each probe shifted the same complex from nuclear extracts obtained from several cell lines, which complies with the ubiquity previously observed for the 3β1-A protein, we conclude that the 3β1-A protein forms the specific band observed with 3β1 and 3β2 oligomers. Thus, the HSD3B2 gene intron 1 contains an imperfectly conserved 3β1-A element that is still able to bind the 3β1-A protein, or the 3β1-A protein only binds to a strictly conserved nucleotide segment within this cis-element.

Sp3 binds the TG boxes of type I and type II intron 1 with different efficiency

The TG box overlapping with the 3β1-A element is conserved between type I and type II introns; however, the previous experiment failed to show significant Sp1 binding to both 3β1 and 3β2 probe, despite the fact that in the earlier study a weak but more conspicuous Sp1 binding was observed (Guerin et al. 1995). While the previous study added at both extremities the unrelated sequence gate, our probe consisted exclusively of nucleotides matching to the region of interest because of a different labeling strategy. Therefore, we reasoned the cause of our undetected Sp1 binding could be the shorter sequence provided at the 3’ end of our probe and that a probe encompassing more of the TG box 3’ flanking region would provide a better spatial requirement for binding of Sp1 family members. For both type I and type II intron sequences, we designed an oligo encompassing fewer 5’ flanking
A

**SW-13**

**LNCaP**

<table>
<thead>
<tr>
<th>Probe:</th>
<th>3β1</th>
<th>3β1m</th>
<th>3β2</th>
<th>3β2m</th>
</tr>
</thead>
<tbody>
<tr>
<td>Competitor:</td>
<td>self</td>
<td>Sp1</td>
<td>self</td>
<td>Sp1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Probe:</th>
<th>3β1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Competitor:</td>
<td>EDTA</td>
</tr>
</tbody>
</table>

B

**Probe:**

<table>
<thead>
<tr>
<th>3β1S</th>
<th>3β2S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antibody:</td>
<td>αSp1</td>
</tr>
<tr>
<td>Competitor:</td>
<td>self</td>
</tr>
</tbody>
</table>

C

**3β1-A element**

| 3β1 | GGACACAGAATGTGCAAAAAAAATGGGAGGAGGAA |
| 3β1m | GGACACAGAATGTGCAAAAAAAATGAAAAGAGGAA |
| 3β1S | TTTGCAAAAAAATGGGAGGAAATAATGAGGCA |
| 3β2 | GGTCATGGAATTTTTGTAAAAATGAGGAA |
| 3β2m | GGTCATGGAATTTTTGTAAAAATGAGGAA |
| 3β2S | TTTTGTAAAAATGAGGAA |
| Sp1 | ATTCGATCGGGGCGGGCGAGC |
nucleotides and more of the 3' flanking region when compared with the 3β1-A element of both genes while maintaining the same length of 36 base pairs (3β1S and 3β2S, shown in Fig. 2C). Despite lacking a few nucleotides of the 3β1-A element 5' prime, these oligos were still able to form the 3β1-A band in gel shift experiments (Fig. 2B, lanes 1, 11 and 13). In addition, they formed three or four new major complexes of slower electrophoretic mobility, numbered in Fig. 2B after assigning number 1 to the 3β1-A band. This result matches the four bands previously reported (Guerin et al. 1995), with the difference that the slower mobility band formed by the type I intron 1 specific probe was remarkably more intense, and comparable with the 3β1-A band (Fig. 2B, lanes 1 and 11). This slower band appeared to be a doublet (named band 4a and 4b) and was efficiently competed by 100-fold molar excess of an Sp1 consensus containing oligo-nucleotide (Fig. 2B, lane 4), which also competed for band 2. As expected, oligos 3β1 and 3β2 only partially competed for band 4 and 2, while efficiently competed out band 1 (Fig. 2B, lanes 5 and 6). Oligo 3β1Sm, which bears the same substitution G4 to As earlier analyzed in Fig. 2A, not only did not compete for complex 1 as expected, but also for complex 2 and 4 (Fig. 2B, lane 7), indicating that both Sp1-related and 3β1-A binding activities rely on the G4 stretch for optimal binding. These mutant oligos did compete the weak complex forming band 3 of intermediate mobility between the Sp1 consensus competed bands. Such competition suggests a protein binding to the 3β1-A element in a region not involving the four Gs forms the complex. The use of antibodies against Sp1, Sp2 and Sp3 revealed that Sp3 accounts for two of the three bands competed by Sp1 consensus. In fact, addition of anti-Sp3 antibody in the binding reactions (Fig. 2B, lane 10) resulted in a significant supershift, as well as in the disappearance of band 4b, the lower and more conspicuous band of the doublet, and band 2 likely formed by the smaller Sp3 (73 kDa) alternative translation product (Kennett et al. 1997, Suske 1999). Addition of anti-Sp3 antibody also resulted in a darker 4a band, suggesting that Sp3 has the strongest affinity for binding to the region, since only when Sp3 is prevented from binding other factors can access the site. Addition of anti-Sp1 antibody resulted in disappearance of the fainter upper band (4a) of the doublet and a weak supershift could occasionally be seen (Fig. 2B, lane 8). Addition of anti-Sp2 antibody had the same effect (Fig. 2B, lane 9), with the difference that a slightly more intense supershifted band was seen, with mobility lower than Sp1 supershift. Surprisingly, the type II intron 1 TG box bound Sp3 less efficiently, as demonstrated by less Sp3 shift when 3β2S oligo was used as a probe (Fig. 2B, compare lane 13 with lane 11), and also by its partial ability to compete the doublet formed with labeled 3β1S (Fig. 2B, lane 3). By densitometry analysis we determined that Sp3 binds to the type II TG box 3-fold less than the type I TG box (data not shown). Sp3 binding activity could be brought to the same extent as seen for the type I intron 1 when an oligo bearing a substitution A+136 to G (Fig. 2C) was used (data not shown), which reproduces the 3' flanking region of the type I TG box. Despite the fact that Sp2 antibody did result in disappearance of the faster band of the doublet, we cannot conclude that Sp2 is part of the observed complex.

Figure 2 Identification of a 3β1-A element in the HSD3B2 gene intron 1. (A) Left: EMSA result obtained when 3 μg nuclear extract from the indicated cell lines were incubated with 3β2 probe with and without self-competition with 25-fold molar excess of unlabeled oligo, as described in Materials and methods. Middle: EMSA result obtained when 3 μg nuclear extract from SW-13 cells were incubated with probes specific to either type I (3β1) or type II (3β2) intron, with and without 100-fold molar excess of the indicated competitor oligonucleotides, as described in Materials and methods. --, indicates no competition. Right: 5 μg nuclear extracts from LNCaP cells were incubated with 3β1 probe in the presence of EDTA 10 mM, with and without 5 mM zinc chloride. The competition with 100-fold molar excess of cold 3β2, 3β2m and Sp1 oligomers is also shown. (B) Left: EMSA result obtained when 3 μg nuclear extract from SW-13 cells were incubated with oligo 3β1S. Competitor oligos were used at 100-fold molar excess. Middle: the same amount of nuclear extracts as above was pre-incubated with anti-Sp1, anti-Sp2 and anti-Sp3 antibodies as described in Materials and methods. Right: a different EMSA experiment compares complexes formed by 3β1S and 3β2S probes with 3 μg nuclear extract from SW-13 cells. (C) The upper nucleotide sequence of the oligomers employed in these gel shift analysis is shown. Oligos are aligned to center the shared regions encompassing the 3β1-A element, evidenced by a line above the 3β1 oligo sequence. The outlined nucleotide in 3β2S and 3β1S identifies the single nucleotide mismatch at their 3' end.

www.endocrinology.org
as anti-Sp2 antibody was able to supershift Sp1 protein expressed in Drosophila Schneider’s cells (not shown), which do not express Sp1 family members, thus indicating cross-reactivity. No other cross-reactivity was observed in this study.

Mapping the boundaries between 3β1-A and Sp3 binding activities

Our EMSA results indicated that 3β1-A and Sp3 binding to the 3β1-A element and TG box respectively are mutually exclusive, as no lower mobility complex indicative of a protein–protein interaction could be seen. Therefore, we were interested in creating mutants able to bind either the 3β1-A protein or Sp3 to be able to discern the function of each of those two factors in relation to the HSD3B2 gene basal promoter activity upon transfection experiments. To precisely define how and at what extent the G4 stretch is shared between transfection experiments. To precisely define how and at what extent the G4 stretch is shared between the apparently overlapping cis-elements, we compared 3β2S with two mutant oligonucleotides. Oligo M1 changed the last two Gs of the G4 stretch to As, while oligo M2 changed the first two Gs in the same way (Fig. 3B). When used in EMSAs, M1 was unable to bind Sp3, while the 3β1-A band was unaffected (Fig. 3A, left, lane 2). M2 bound Sp3 similarly to as is seen with wild type 3β2S, while the 3β1-A band was severely diminished (Fig. 3A, left, lane 3). This result indicates that the distal two Gs in the sequence 5′-TGGGGT-3′ are essential for Sp3 binding, and not necessary for 3β1-A binding. The proximal two Gs are not required for Sp3 binding, but are important for 3β1-A factor binding.

Next we asked what is the contribution of the long stretch of As to 3β1-A factor binding, which accounts for as many as eight and six residues in the HSD3B1 and HSD3B2 gene respectively. We disrupted the A6 sequence by changing the five As immediately preceding the TG box to CCTTC (oligo 3β2mS, Fig. 3B) and found that the substitution completely prevents formation of the 3β1-A band (Fig. 3A, lane 4) while unable to form any additional complex. To determine what is the minimal number of As required for restoring the 3β1-A binding activity, we created a series of three oligos that reinsert two, three and four As respectively (A2, A3, and A4, Fig. 3B). A2 and A3 that restore two and three of the As in front of the TG box, were unable to support 3β1-A protein binding (Fig. 3A, lanes 5 and 6), while addition of four As resulted in a fully restored 3β1-A complex (Fig. 3A, lane 7). We also tested a mutant oligo bearing a deletion of the TGGGG sequence, ΔT(G)4. This deletion, which was expected to prevent binding of the 3β1-A and Sp3 proteins, surprisingly abrogated formation of every band seen with the wild type probe except for the 3β1-A band (Fig. 3A, lane 8). In addition to the 3β1-A band, M2 and all oligos bearing mutations in the A stretch also affected formation of a faster band that appeared to be as specific as the 3β1-A band (Fig. 3A, appearing in lanes 1, 2 and 8), but whose intensity and/or formation was not consistent in all our EMSA experiments, as later discussed. Lack of competition with the 3β1-A protein is likely the reason for the stronger Sp3 binding observed with B2mS, A2 and A3, as well as with the unknown protein that forms band 3, the complex of intermediate mobility between the slower and faster Sp3 band, whose binding to the probe is dependent on the integrity of the A6 stretch (Fig. 3A, lanes 5–8 all lack this complex). However, this explanation does not apply to oligo A4, which is expected to show less Sp3 binding as a consequence of its ability in restoring the 3β1-A band. Interestingly, this oligo, despite binding the 3β1-A factor, did not form the faster band discussed above, which is seen with all other samples forming the 3β1-A complex (Fig. 3A lanes 1, 2 and 9). It is possible that, in addition to preventing the formation of band 3, the changes in the A stretch carried by A4 also affect the ability to form the faster complex, which apparently results in an increased Sp3 affinity for the TG box. However, conformational changes as a consequence of the alteration in the A6 stretch carried by B2mS, A2, A3 and A4, may also explain the stronger affinity for Sp1 family members of these mutated oligo, independently from their ability to bind or not the 3β1-A protein or the faster band.

The results of the mutational analysis presented herein confirmed the hypothesis that the 3β1-A protein only recognizes a short segment of the 24 bp element. Indeed it was evident that the conserved sequence AAAATGG represents the core-binding region for the 3β1-A factor, and that the Sp3 core-binding region is the sequence GGTGG. Therefore, the 3β1-A element and the TG box of the HSD3B2 and HSD3B1 gene appear to be flanking rather than overlapping cis-elements.
The discovery that the 3β1-A protein only requires two of the four Gs to bind the 3β1-A element, explains the apparently aberrant behavior of oligo 3β2S. Middle: the effect of mutations altering the stretch of As in the type II 3β1-A element is shown. Right: the complexes formed by oligo ΔTG₄ are shown. (B) Nucleotide sequence of the 3β2S oligonucleotide and its mutants used in this study. Small letters indicate the mutated nucleotides. The name of each oligo is shown on the left side of each sequence. The enlarged sequence at the bottom of the oligo list summarizes the analysis results, with the minimal nucleotides required for 3β1-A factor binding showed in bold capital letters, and the minimal nucleotides required for Sp3 binding showed in cursive capital letters. Dotted lines substitute for the flanking nucleotides not involved in direct binding. The ability of each oligo to bind YY1 and/or Sp3 is summarized in the panel at the right side of the sequence alignment. +, binding; −, no binding; −/+ , partial binding.

Figure 3 Dissection of the 3β1-A element and TG box in the HSD3B2 gene intron 1. EMSA results obtained with 5 µg nuclear extract from SW-13 cells and each of the labeled oligonucleotides whose sequence is shown in (B). (A) Left: comparison of the binding ability of mutants M1 and M2 with the shifted band pattern formed by oligo 3β2S. Middle: the effect of mutations altering the stretch of As in the type II 3β1-A element is shown. Right: the complexes formed by oligo ΔTG₄ are shown. (B) Nucleotide sequence of the 3β2S oligonucleotide and its mutants used in this study. Small letters indicate the mutated nucleotides. The name of each oligo is shown on the left side of each sequence. The enlarged sequence at the bottom of the oligo list summarizes the analysis results, with the minimal nucleotides required for 3β1-A factor binding showed in bold capital letters, and the minimal nucleotides required for Sp3 binding showed in cursive capital letters. Dotted lines substitute for the flanking nucleotides not involved in direct binding. The ability of each oligo to bind YY1 and/or Sp3 is summarized in the panel at the right side of the sequence alignment. +, binding; −, no binding; −/+ , partial binding.

The discovery that the 3β1-A protein only requires two of the four Gs to bind the 3β1-A element, explains the apparently aberrant behavior of oligo ΔTG₄ that formed the 3β1-A band despite bearing a deletion encompassing the G₄ stretch. Analysis of this oligonucleotide sequence reveals that the two G residues critical for 3β1-A protein binding are provided by the sequence immediately following the deletion, thus restoring the AAAATGG sequence (Fig. 3B). Instead, Sp3 binding is no longer supported by this oligo, indicating that the sequence GGAGG is unable to functionally replace the sequence GGTGG. The effect of each mutation on the ability to sustain 3β1-A and/or Sp3 protein binding is summarized in Fig. 3B in the right panel next to each oligo sequence.
YY1 is the protein forming the 3β1-A complex

The sequence AAAATGGG, that represents the core-binding region of the 3β1-A factor, is almost identical to the UCR (upstream conserved region) of the Moloney murine leukemia virus, which is known to bind the transcription factor UCRBP (Flanagan et al. 1992), later identified as the murine YY1 (Safrany & Perry 1993). The UCR, which is only one of many consensus sites known to bind YY1 (Shi et al. 1997), all sharing the CCAT core motif, was later found also in many other viral and cellular genes. The 3β1-A factor core-binding region matches the UCR nucleotide sequence seven out of nine nucleotides (Fig. 4C). To find out if the 3β1-A band is formed by YY1, an oligonucleotide containing the UCR element was tested in EMSA experiments in competition with
the 3β2S probe. As expected, a 100-fold molar excess of unlabeled 3β2S oligo competed any band formed by the probe with nuclear extract from three different cell lines (Fig. 4A, lanes 1 and 2, 4 and 5 and 7 and 8). Addition of 100-fold molar excess of cold UCR oligo resulted in the disappearance of the 3β1-A band (Fig. 4A, lanes 3, 6 and 9), suggesting that the factor binding to the 3β2S probe – and therefore 3β1S as well as 3β1 and 3β2 – is able to recognize the UCR element. The competition was specific, as the Sp3 band was unaffected. Furthermore, after incubation with SW-13 cell nuclear extract 32P-labeled UCR was shifted into a major complex with the same electrophoretic mobility as the 3β1-A band formed by 3β2S probe (Fig. 4B, compare lane 2 with lane 4). The complex was specific, as competed out by a 100-fold molar excess of unlabeled UCR (Fig. 4B, lane 5). Addition of a rabbit polyclonal antibody directed against full-length YY1 into the binding reactions resulted in the disappearance of both the 3β1-A band formed by 3β2S probe and the specific complex formed by UCR probe, and a complicated pattern of supershifted bands was seen with both probes (Fig. 4B, lanes 1 and 3). These results suggest that the protein forming the 3β1-A band is the transcription factor YY1 or a factor antigenically related to YY1. As YY1 is known as a 68 kDa protein, while the size of the 3β1-A factor was estimated to be 37 kDa by Guerin et al. (1995), we compared the size of the band shifted by labeled UCR oligomer from SW-13 cell nuclear extract with in vitro-translated YY1. The in vitro obtained protein from plasmid pCMX-YY1 in gel retardation assays did shift the UCR probe forming a complex of the same mobility as the one observed with nuclear extracts (Fig. 5A, compare lanes 1 and 3). Such a band was specific to UCR (Fig. 5A, lane 4), and it was not formed when the probe was reacted with an in vitro transcription/translation reaction performed in absence of template (Fig. 5A, lane 2), which indicates that the band is not formed by interaction of UCR with a protein from the rabbit reticulocyte lysate. Western blotting analysis revealed that anti-YY1 antibody – the same antibody used in our EMSAs – recognized a band of molecular mass slightly above 55 kDa for the in vitro transcribed/translated YY1 protein (Fig. 5B, lanes 1–7).
lane 2) and nuclear extracts from both SW-13 (Fig. 5B, lane 3) and LNCaP cells (Fig. 5B, lane 6) as well. This band was not formed in negative control reactions consisting of an *in vitro* transcription/translation coupled reaction without template (Fig. 5B, lane 1). This result indicated that the 3β1-A band formed by nuclear extracts with UCR, as well as 3β2 and 3β2S oligomers, corresponds to full-length YY1 protein. In addition to the full-length band, anti-YY1 antibody recognized also two additional bands of lower molecular mass. A band approximately of 35 kDa and that it can eventually be co-purified from the lysate. To confirm that YY1 binds the 3β2 probe and that it can eventually be co-purified from the bound proteins as differently sized polypeptides, we have used a modification of the Acrydite technology method for the isolation of DNA binding proteins recently proposed by Nelson et al. (2002). Briefly, proteins from SW-13 or LNCaP cell nuclear extracts were incubated in binding buffer with a 3β2 oligomer whose upper strand carried an Acrydite moiety at its 5′ prime. Then the oligomer was immobilized to acrylamide as previously described (Nelson et al. 2002) and the unbound proteins were separated from the bound complexes by electrophoresis. An aliquot of a crude preparation, consisting of eluted and concentrated 3β2-binding proteins from recovered acrylamide slices, reacted with anti-YY1 antibody in Western blots, forming three bands of the same molecular mass as obtained with unbound nuclear extract proteins (Fig. 5B, lane 5). Likely because of contaminants in these samples and our loss-prone modification of the original method, the antibody-recognized bands were less intense than expected, particularly when compared with the bands from unbound nuclear proteins. Coomassie staining of an SDS gel of the same preparation revealed bands whose size were identical to the YY1 specific bands (Fig. 5C). In addition, an intense band of molecular mass intermediate to 35 and 49 kDa was also observed, that was not recognized by anti-YY1 antibody. Sp3 antibody, also a rabbit polyclonal antiserum, recognized the proper Sp3 specific bands of 105 and 73 kDa, but failed to recognize the 30 and 35 kDa bands (Fig. 5B, lane 8). These results indicate that at least three different YY1 polypeptides were specifically recognized by anti-YY1 antibody, corresponding to the full-length protein, and to two fragments of about 35 and 30 kDa. They also indicate that these YY1 polypeptides bind the 3β2 oligomer, which is consistent with the previously estimated size of the 3β1-A factor and is likely due to proteolytic degradation of YY1.

**Existence of a second YY1 site within the HSD3B2 and HSD3B1 gene intron 1**

Established that the sequence AAAATGG within the 3β1-A element binds YY1, as well as that YY1 and the 3β1-A factor are identical, we observed that a similar sequence is located in reverse orientation about 35 nucleotides downstream from the TG box in type II intron 1. Because of a substitution T/G, the site is not perfectly conserved in the type I intron 1 at the same location. However, another substitution A/G at nucleotide +138 creates a variant of the sequence immediately following the TG box in the type I intron (AAAATGAGG in type I, AAAATAAGG in type II). To verify whether or not these sequences are able to bind YY1, we designed the oligos shown in Fig. 6B and tested them in EMSAs. As shown in Fig. 6A, the EMSA experiments confirmed the ability of YY1 to bind the HSD3B2 gene intron 1 35 bp downstream from the 3β1-A element (lane 5). YY1 recognized oligo 147 with an affinity slightly lower than the efficiency exhibited by binding to the UCR and even to the 3β1-A element. This discovery is in agreement with our findings that only the two proximal G residues are absolutely required for YY1 binding within the 3β1-A element (Fig. 3). Further in accord with the results of the mutational analysis shown in Fig. 3, the corresponding sequence in type I intron 1 was not able to bind YY1 (Fig. 6A, lane 6) because of its bearing only three out of the four required As. Oligo 126, matching the putative YY1 binding site created by the A+138/G substitution was able to bind YY1 (Fig. 6A, lane 8), but with much lower affinity as only one of the two required Gs is present. Oligo 125, matching the corresponding region in the type II intron was unable to bind YY1 as it lacks the two required proximal Gs (Fig. 6A, lane 7). In Fig. 6B, along with the double stranded sequence of the oligomers used in this study, the ability of each oligo in binding YY1 is exemplified next to each sequence by a positive (binding) or negative (no binding) sign.
Figure 6  Existence of a second YY1 binding site within the HSD3B1 and HSD3B2 gene Intron 1. (A) EMSA experiment carried out with 5 µg SW-13 nuclear extracts incubated with each of the oligomers indicated as probe. Competition with 100-fold molar excess of UCR and Sp1 oligomers is also shown for the 3β2S probe. −, no competition. An arrow points at the YY1 band, previously known as 3β1-A band. Only the slower migrating 3β2S probe is shown. (B) The double stranded sequence of the oligomers used in the study is shown. The name of the mutant oligomers specifies the first and last nucleotide of the intron 1 sequence they encompass. (I), (II) identifies oligos matching to type I or type II introns respectively. The nucleotides involved in the YY1 binding core within the 3β1-A element is shown as a reference, with dotted lines replacing the 5' and 3' flanking nucleotides. The black arrow indicates orientation of the YY1 binding site. Dotted lines encase the putative YY1-like binding region in each oligomer for sequence comparison. The ability of each oligomer to bind YY1 is summarized. ++, strong binding; +, binding; +/-, weak binding; −, no binding. (C) Upper: a cartoon illustrates the different organization of the two YY1 binding sites in respect of the TG box in the HSD3B1 and HSD3B2 gene intron 1. Solid and dashed boxes represent stronger and weaker YY1 binding respectively. Open or filled square boxes represent weaker and stronger Sp3 binding to the TG box respectively. A narrow rectangular box represents the 24 nucleotides previously identified as the 3β1-A element in the HSD3B1 gene intron 1. A similar schematization is used to represent the not perfectly conserved region in the HSD3B2 gene. Lower: the sequence alignment of the type I and type II intron 1 from nucleotide +111 to +174 (type I) or to +173 (type II) shows the nucleotide changes involved in the formation of the second and not conserved YY1 binding site. Small letters identify the nucleotide substitutions, dotted lines replace missing nucleotides. Bold letters outline the nucleotide essential for Sp3 binding. Arrowheads identify the nucleotides involved in YY1 binding and the orientation of such binding. Bold solid arrowheads indicate strong binding, a bold open line arrowhead indicates weaker YY1 binding.
This analysis revealed that YY1 recognizes the same sequence in two different regions within intron 1 of both the HSD3B1 and HSD3B2 genes. One site is conserved between the two genes for nucleotide sequence and location, it being the seven conserved nucleotides of the 3β1-A element immediately juxtaposed to the 5′ prime of the TG box; the second site differs between the two genes for location and orientation. In type I intron 1 the second YY1 site is a slight variant of the AAAA TGG sequence located eight nucleotides downstream from the 3β1-A element, immediately following the 3′ prime of the TG box. In type II intron 1 the second site lays in reverse orientation 35 bp apart from the 3β1-A element. A schematic representation of these findings is presented in Fig. 6C.

**YY1 enhances the HSD3B2 gene basal promoter activity by binding to the two identified binding regions within the gene intron 1**

To explore whether or not YY1 is functionally involved in the boost of basal promoter activity mediated by the HSD3B2 gene intron 1, we overexpressed YY1 in the presence of the same reporter constructs (shown Fig. 1B) we earlier used to assess the functional relevance of the intron in transient transfection assays. As shown in Fig. 7A, co-transfection with increasing amounts of YY1 expression plasmid stimulated the full-length construct pGL3-HSD2(−517/+193) in a dose-dependent manner up to 3-fold, while it failed to stimulate the intron-less construct.
pGL3-HSD2(−517/+23). This result showed that YY1 stimulates the luciferase expression driven by the HSD3B2 gene sequence by binding within intron 1. To confirm YY1 as the factor responsible for the enhanced basal transcription observed in the presence of intron 1 we investigated the contribution of each of the two YY1 binding sites by comparing the activity of a luciferase reporter construct encompassing the HSD3B2 gene sequence from nucleotide −246 to nucleotide +193, pGL3-HSD2(−246/+193), with reporter constructs in which the same region was abrogated of YY1 binding to either the 3β1-A element or the distal site or both. For this purpose, the four As required for YY1 binding were changed by site-directed mutagenesis using oligo B2mS (Fig. 3) to abolish YY1 binding in front of the TG box; an oligo identical to oligo 147 (shown in Fig. 6), except for having the four critical As changed to CCTT, was used to eliminate YY1 binding to the distal site. We also tested a construct bearing the ΔT(G)₄ deletion that eliminates the sequence TGGGG (Fig. 3), thus preventing Sp3 binding but not YY1 binding. Fig. 7B presents the results of transient transfections in SW-13 cells obtained with pGL3-HSD2(−246/+193), with reporter constructs in which the same region was abrogated of YY1 binding to either the 3β1-A element or the distal site or both. For this purpose, the four As required for YY1 binding were changed by site-directed mutagenesis using oligo B2mS (Fig. 3) to abolish YY1 binding in front of the TG box; an oligo identical to oligo 147 (shown in Fig. 6), except for having the four critical As changed to CCTT, was used to eliminate YY1 binding to the distal site. We also tested a construct bearing the ΔT(G)₄ deletion that eliminates the sequence TGGGG (Fig. 3), thus preventing Sp3 binding but not YY1 binding. Fig. 7B presents the results of transient transfections in SW-13 cells obtained with pGL3-HSD2(−246/+193) and its mutant variants, which did not involve co-transfection of YY1 expression vector. Preliminary experiments revealed that the basal activity of pGL3-HSD2(−246/+193) was comparable with the basal activity of pGL3-HSD2(−517/+193), indicating that the HSD3B2 gene sequence from −517 to −246 does not significantly affect the basal promoter activity. Surprisingly, inability of YY1 to bind its site within the 3β1-A element had little effect on the basal promoter activity (Fig. 7B, activity bar 2). A similar effect was seen with the construct that bears a deletion of the TG box (Fig. 7B, activity bar 3). Therefore, neither YY1 nor Sp3 binding to the juxtaposed 3β1-A element and TG box plays any major role in the intron 1-mediated increase of the HSD3B2 gene basal promoter activity. Instead, preventing binding of YY1 to the distal site resulted in a 50% decrease of basal transcription (Fig. 7B, activity bar 4). Complete abrogation of YY1 binding within intron 1 resulted in a basal promoter activity that was only 30% of the wild type activity (Fig. 7B, activity bar 5). Remarkably, this last result reproduced the loss of activity seen with the promoter construct missing the whole intronic sequence and also it made evident that Sp3, or any other Sp1 family members, has no ability to direct the boost in basal activity through intron 1 of the HSD3B2 gene. Thus, the proximal YY1 site within the 3β1-A element and the distal YY1 binding site contribute to boost the HSD3B2 gene basal transcription. Similar results were obtained with COS-7 and LNCaP cells (data not shown).

Discussion

Despite the pivotal role of the 3β-HSD type II enzyme in the biosynthesis of all steroids, the molecular mechanisms that regulate the gene expression are yet to be fully understood. Lately, important studies have been published concerning the significance of SF-1 in the HSD3B2 gene expression and its synergistic effect with phorbol ester (Leers-Sucheta et al. 1997), as well as its regulation by cytokines (Cote et al. 2000, Gingras et al. 2001), gonadotropins (Feltus et al. 1999) and glucocorticoids (Feltus et al. 2002). However, no study has investigated the role of intron 1 in the HSD3B2 gene basal activity so far, despite a previous finding that the highly homologous HSD3B1 gene requires the activating 3β1-A element within the first intron to keep the basal promoter activity to its full potential by binding the unknown 3β1-A factor. In this study we have shown that the HSD3B2 gene intron 1 is required to achieve maximal basal activity levels of the promoter region. Despite the suggestion that the HSD3B1 and HSD3B2 genes may not share the 3β1-A element (Guerin et al. 1995), we have proved the existence of such a cis-acting element in the intron 1 of the HSD3B2 gene. More specifically, we have proved that the 3β1-A protein binds to a conserved region of seven nucleotides of the 3β1-A element in both genes, through a mutational analysis that has ultimately helped us pinpoint similarities between the 3β1-A binding core and the UCR core recognized by the multi-role transcription factor YY1. We have shown evidence that YY1 is indeed the protein binding in vitro to the 3β1-A element by competition analysis and use of anti-YY1 antibody.

YY1 complies with many of the characteristics attributed to the 3β1-A factor, as it is ubiquitously and constitutively expressed (Austen et al. 1997, Shi et al. 1997). It contains four zinc-finger domains
with homology to the GLI-Kruppel family of proteins (Shi et al. 1991), thus explaining the 3β1-A protein requirement for zinc ions. However, despite a predicted molecular mass of 44 kDa, YY1 is known to migrate as a 65–68 kDa protein in SDS gels (Shi et al. 1997), whereas the 3β1-A factor was attributed a molecular mass of 37 kDa by UV-crosslinking analysis (Guerin et al. 1995). This apparent discrepancy can be explained by recent findings that claim YY1 proteolytic mechanisms as part of degradation processes and even tissue-specific signal transduction pathways. Specifically, YY1 was shown to be the target of calpein II in a mechanism aimed at downregulating YY1 protein during muscle development (Walowitz et al. 1998), which generates a 40 kDa polypeptide. Moreover, a nuclear cathepsin B-like protease activity appears to degrade YY1 (Pizzorno 2001) generating two fragments of about 30 and 40 kDa, a phenomenon associated with the progression of undifferentiated to differentiated NT2 cells upon treatment with retinoic acid. In this case, the larger cleavage product, representing the carboxy-terminal portion of YY1 containing the zinc-finger domain, was shown to bind the cognate DNA consensus forming a faster complex in gel shift assays. Bovolenta et al. (1999) found that administration of interleukin 2 to individuals affected by HIV-induced downregulation of YY1 protein through activation of an unidentified proteolytic activity, which resulted in appearance of a faster YY1 band of unknown molecular mass in gel shift analysis. Interestingly, when YY1 was first isolated for its ability to bind the UCR site, a 40 kDa antigenically related protein was co-purified with it (Becker et al. 1994) and found to account for much of the UCR binding activity in T-lymphocytes along with YY1. In light of the latest findings, the 40 kDa protein could have been a truncated form of YY1. Therefore, it cannot be excluded that in the previous study on the HSD3B1 gene (Guerin et al. 1995) during the procedures required for UV-crosslinking analysis a pre-existing truncated YY1 fragment was also isolated. In addition to the 37 kDa band attributed to the 3β1-A protein, according to Guerin and coworkers, the UV-crosslinking analysis produced two additional specific complexes that, once approximately corrected for the contribution of the oligonucleotide, did yield molecular mass compatible with full-length YY1. In support of the above suggestions, we did observe frequently but not constantly appearance of a band of faster mobility with respect to the YY1/3β1-A band in our gel shift experiments, specific to both 3β1S/3β2S and UCR probes (Fig. 3, and seen also in Fig. 6). Formation of this faster band seems to be sensitive to slight experimental changes of pH or ionic strength that facilitate binding of pre-existing factors/fragments, and also it might be influenced by the nucleotides flanking the consensus, as our deletion construct ΔT(G)4 was able to form this faster complex at all times. In addition, use of anti-YY1 antibody in gel shift experiment showed more than one super-shifted band. We exclude that YY1 fragmentation occurred during the manipulation of the nuclear extracts and incubation of the binding reactions. Indeed, despite absence of protease inhibitors in our nuclear extracts, incubation of the extract at 37 °C for 1 h before probe addition in EMSA binding reactions did not favor the appearance of faster bands in control experiments that did not show these bands in the untreated samples (not shown). Although we cannot at this time explain their origin, neither their significance in relation to the HSD3B2 gene basal activity, we have demonstrated that two bands of similar molecular mass could be isolated from nuclear factors bound to the 3β1-A element. The 35 kDa YY1 fragment could likely account for the 37 kDa band previously attributed to the 3β1-A factor.

In addition to having characterized YY1 as the factor binding the 3β1-A element, we have identified a second YY1 binding site in both type I and type II introns. Indeed, a sequence 100% identical to the region we have found to bind YY1 within the 3β1-A element was found to bind YY1 at the end of the HSD3B2 intron 1. Interestingly, this sequence is reversed in respect of the orientation of the identical sequence within the 3β1-A element. A weaker YY1 binding site was found at the 3′ end of the TG box in the HSD3B1 intron 1. This last finding might explain why Sp3 binds with better efficiency the 3β1S probe than the 3β2S probe. We could bring the band intensity formed by Sp3 bound to 3β2S to that seen with the 3β1S probe when we changed the only nucleotide that differs between the type I and type II sequence. This one nucleotide mismatch in position +136, A in type II, +138 G in type I (Fig. 2C), does not affect the region strictly required for Sp3 binding,
neither is it part of a region that may represent another Sp1/Sp3 binding site, as demonstrated by the fact that oligos 125 and 126 are unable to shift Sp1 family members (Fig. 6A). It is possible that a G instead of A at the 3′ end of the intron 1 TG box may facilitate Sp3 binding to a region rich in As and Ts, by reducing or disrupting any intrinsic DNA structure. However, a more likely explanation is that in the fewer cases when YY1 is sitting onto the less efficient site outside the 3β1-A element, Sp3 displaces it more easily than it can displace YY1 bound to the stronger site within the 3β1-A element, which results in a better competition of Sp3 for the type I TG box. Our findings suggest the effectors playing a role in the HSD3B1 gene basal activity need to be further investigated, as the conclusions of the previous study were based on the analysis of a mutant that affected not only YY1 binding to the 3β1-A element but also Sp3. While the importance of Sp1 was disregarded in the previous study because of the weaker binding of this factor to the 3β1-A element flanking TG box, we have herein proved the oligomer used in that study could not support proper Sp1 family member binding, and that Sp3 is the Sp1 family member that electively binds the region with intensity comparable with the complex formed by YY1. Nevertheless our results in mammalian cells did not evidence any role for Sp3 in the maintenance of the HSD3B2 gene basal activity, the different organization of the two YY1 binding sites in respect of the TG box in the HSD3B1 gene intron 1, and the consequent improved Sp3 competition for its site in in vitro experiments, suggest that Sp3 may have a different outcome in the HSD3B1 gene basal promoter activity. The importance of the weaker YY1 binding site immediately flanking the TG box at the 3′ end also needs to be assessed.

We have functionally characterized the importance of YY1 in the intron 1-mediated enhancement of the HSD3B2 gene basal promoter activity. Dissimilarly from the HSD3B1 gene (Guerin et al. 1995), disruption of the YY1 binding site within the 3β1-A element did not result in a 3- to 6-fold reduction of promoter activity, an outcome that further contributes to make evident differences in the mechanism implicated in the function of type I and type II intron 1. However, disrupting YY1 binding to the distal site with opposite orientation resulted in a 50% decrease in basal promoter activity. When both sites were simultaneously eliminated, the residual promoter activity was only 30% of the wild type promoter activity, a result functionally identical to the lack of the whole intronic sequence. As the two 35 bp apart YY1 binding regions have opposite orientation, and YY1 is known to induce DNA bending by binding to its cognate sequence (Natesan & Gilman 1993, Kim & Shapiro 1996), one could speculate that YY1 binding to the two regions within the HSD3B2 intron 1 would induce formation of a loop which might help binding and/or interaction of other factors involved in the HSD3B2 gene basal transcription. As it is known that TBP and TFIID require bending of the TATA box region for optimal binding (Parvin et al. 1995), it is tempting to hypothesize that YY1 binding to intron 1 would help maintain a DNA conformation that facilitates the RNA polymerase complex formation onto the imperfect TATA box (Lachance et al. 1991) of the HSD3B2 gene. Similarly to our study, Kerr et al. (1997) found that extending the human bone sialoprotein gene promoter to include the first 68 nucleotides of intron 1 resulted in a 3- to 7-fold increase in reporter activity in UMR106–01 BPS cells. Interestingly, the human bone sialoprotein gene possesses an inverted TATA box, and one YY1 binding site was found within the first 30 bp of intron 1. The explanation we suggest is more likely to describe the mechanism of YY1 action through its binding sites in the HSD3B2 gene intron 1 rather than YY1 acting by changing the DNA structure and/or chromatin organization by its ability to recruit factors with intrinsic histone acetylase activity (Thomas & Seto 1999). Indeed, in preliminary transfection experiments, 66 nM trichostatin A, a drug known to inhibit histone deacetylase thus boosting gene activity by favoring an open chromatin structure, did increase the basal activity of the wild type and double YY1 deletion constructs to a similar extent (3-fold) in transfection experiments. This suggests that YY1 is not functioning at the intron 1 level by favoring histone acetylation.

Taken together our results indicate that the HSD3B2 gene intron 1 is required for maintaining the gene basal promoter activity at the highest levels observed in transfection experiments with a reporter system, and that YY1 is the factor that mediates such effect by binding to two different regions within intron 1. As we have also identified YY1 to be the 3β1-A protein, the trans-activator
that was attributed a similar role in the HSD3B1 gene, this study points to a major role of YY1 in affecting the basal levels of 3β-HSD activity. Additional studies will be required to explore the mechanisms implicated in YY1 contribution to the basal activity of both the HSD3B1 and HSD3B2 genes.

An important consideration arising from this study and the previous study on the HSD3B1 gene is that nucleotide sequence variants occurring within the HSD3B1 and HSD3B2 gene intron 1 may result in more or less profound changes in gene basal activity, and consequently they should not be excluded from being characterized merely because of their occurrence within an intronic sequence.

Acknowledgements

We thank Dr Yang Shi for allowing us to use pCMV-YY1. We are most grateful to Dr Baruch Frenkel and Dr Amy Lee for critical review of the manuscript. We wish to acknowledge Claudia Wehie for skillful technical assistance, and members of the J K R’s laboratory, Troy Phipps above all, for assistance with the ABI PRISM 3100 DNA Sequencer. This work was supported by Department of Defense grant # 001280 awarded to J K R

References


Becker KG, Jedlicka P, Templeton NS, Liotta L, & Ozato K 1994 Characterization of hUCRBP (YY1, NF-E1, δ); a transcription factor that binds the regulatory regions of many viral and cellular genes. *Gene* **150** 259–266.


Kennett SB, Ulvdad AF & Horowitz JM 1997 Sp3 encodes multiple proteins that differ in their capacity to stimulate or repress transcription. *Nucleic Acids Research* **25** 3110–3117.


Lala DS, Rice DA & Parker KL 1992 Steroidogenic factor I, a key regulator of steroidogenic enzyme expression, is the mouse homolog of fushi tarazu-factor I. *Molecular Endocrinology* **6** 1249–1258.
Enhancement of the HSD3B2 gene basal promoter activity by YY1


Pizzorno MC 2001 Nuclear cathepsin B-like protease cleaves transcription factor YY1 in differentiated cells. Biochimica et Biophysica Acta 1536 31–42.


Shi Y, Seto E, Chang LS & Shenk T 1991 Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and relief of repression by YY1, a human GLI-Kruppel-related protein, and relief of repression by YY1, a human GLI-Kruppel-related protein, and relief of repression by YY1, a human GLI-Kruppel-related protein. Cell 67 377–388.


Thomas MJ & Seto E 1999 Unlocking the mechanisms of transcription factor YY1: are chromatin modifying enzymes the key? Gene 236 197–209.


Received in final form 8 March 2004
Accepted 26 March 2004
Made available online as an Accepted Preprint 7 April 2004