Characterization of the human sex hormone binding globulin (SHBG) gene and demonstration of two transcripts in both liver and testis

S.Gershagen, Å.Lundwall and P.Fernlund

Department of Clinical Chemistry, University of Lund, Malmö General Hospital, S-21401 Malmö, Sweden

Received July 20, 1989; Revised and Accepted October 6, 1989

EMBL accession nos. X16349–X16351(incl.)

ABSTRACT

A genomic cosmid clone for human sex hormone binding globulin (SHBG), a liver-secreted plasma glycoprotein that binds sex steroids, was isolated with a previously characterized liver cDNA as probe. Southern blot analysis of genomic DNA indicated that only one SHBG gene is present in the human haploid genome. A 3.8 Kb Xba I-fragment of the clone containing the entire coding region of SHBG was sequenced. The SHBG gene has 8 exons. The 5'-end preceding the translation start site had no TATA box or CAAT box promoter elements. Screening of a human testis cDNA library resulted in the isolation of two distinct cDNA forms. One cDNA was identical with the previously characterized liver SHBG cDNA, thus suggesting that human SHBG and the androgen binding protein (ABP) produced by Sertoli cells are coded for by the same gene. The second cDNA differed from the first by having exon I exchanged with a completely different sequence and exon VII deleted. An exon coding for the 5'-end of this cDNA was found in the cosmid clone 1.5 kb upstream of the first SHBG exon. Primer extension experiments showed the alternatively spliced transcript corresponding to the second cDNA to be present in both liver and testis. From the primary structure of this putative SHBG-gene-related protein, it may be deduced that it is a protein very different from SHBG and probably without steroid binding activity.

INTRODUCTION

Human sex hormone binding globulin (SHBG) is a plasma glycoprotein that binds several sex steroids with high affinity such as dihydrotestosterone and testosterone (1, 2). SHBG in plasma is a dimer of two essentially identical monomers, the dimer binding only one steroid molecule (2, 3). The primary structure of the human SHBG monomer has been determined both by amino acid sequencing (4) and from liver cDNA sequences (5–7). It is a single peptide of 373 amino acids probably containing three carbohydrate side chains.

Isoelectric focusing has shown human SHBG to be heterogeneous (8–13), a heterogeneity probably reflecting a slight variation in the structure of the monomer, which is demonstratable by SDS-PAGE (10, 12, 13) and amino acid sequencing (4, 12, 13). Furthermore, the monomer sizes define two structural variants of SHBG present in the general population (13). The cause of the SHBG-monomer heterogeneity as well as of the difference between the population variants has not yet been fully clarified, although it may be differences in glycosylation (9, 10, 12).

Human SHBG has been shown to be produced by the liver (14), which is thought to be the source of SHBG in plasma. The Sertoli cells of the human testis produce a protein, generally referred to as androgen binding protein (ABP) (15). In the human the testis protein (ABP) is not so well characterized as the plasma protein (SHBG), though the

two proteins have been shown to have many characteristics in common (16). Although the rat lacks plasma SHBG, its ABP, which is better characterized than the human protein and has been cloned (17), has an amino acid sequence that is 68% identical to that of human SHBG (18). This finding may suggest that human SHBG and human ABP are the same proteins, produced in two different organs but originating from the same gene. To test this we have cloned and sequenced the human SHBG gene and also characterized the gene transcripts in liver and testis.

MATERIALS AND METHODS

Reagents

Radioactive nucleotides were obtained from Amersham (Aylesbury, Bucks, UK). The M13 mp18/mp 19 vectors and oligo (dT)-cellulose were obtained from Pharmacia (Uppsala, Sweden) and Sequenase® was purchased from United States Biochemicals (Cleveland, OH). Biotinylated nylon membranes were supplied by Pall Biosupport Division, Glen Cove, NY. A human testis cDNA library cloned in lambda gt 11 was purchased from Clontech (Palo Alto, CA). All other enzymes were obtained from IBI (New Haven, CT) if not stated otherwise. Oligonucleotides were synthesized by the phosphoramidite method employing an Applied Biosystems model 380A DNA synthesizer (Foster City, CA) and purified by three diethylether extractions followed by ethanol precipitation. The following oligonucleotides were synthesized for use in hybridization and/or primer extension studies: oligonucleotide A, GGCCCATCCCTGGCGGGTGTG (complementary to sequence A in Figure 5); oligonucleotide B, GCACTGCGGGGAGCCTTTGA (complementary to sequence B in Figure 3); and oligonucleotide C, GGGGTTCTTAGGTGGA-GCTTTAATGGG (complementary to sequence C in Figure 5). The cloning of the SHBG cDNA (SHBG I cDNA) used in this work has been described previously (5).

Preparation of radioactive probes

Restriction fragments of SHBG cDNA for use as probes were radioactively labeled with [α -³²P] dCTP (3000 Ci mmol⁻¹) by random primer extension (19). Oligonucleotides were labeled with T₄ polynucleotide kinase and [γ -³²P] ATP (20).

Southern blot analysis

Standard procedures were used for the isolation of human genomic DNA or cosmid DNA, restriction enzyme digestion and agarose gel electrophoresis (20). DNA was transferred from gels to nylon membranes using the conditions recommended by the manufacturer. Membranes were hybridized at 65°C under standard conditions and the membranes were washed to a final salt concentration of 0.2×SSC at 65°C (20). When oligonucleotides were used as hybridization probes, both the hybridization and the washing temperature was 42°C.

Northern blot analysis

Total RNA was isolated from human testis and liver by the method of Chirgwin *et al* (21). Poly (A⁺) RNA was prepared from total RNA by chromatography on oligo (dT)-cellulose (22). Northern blot analysis was performed essentially as described by Thomas (23). The RNA was separated by electrophoresis on formaldehyde agarose gels, blotted onto nylon membrane, using the conditions recommended by the manufacturer, and hybridized overnight with ³²P-labelled SHBG I cDNA dissolved in 50% formamide, 5×SSC, 50 mM sodium phosphate, pH 6.5, 5×Denhardt's solution, 0.5 mg/ml denatured salmon sperm DNA (Sigma, St Louis, MO) and 1% SDS at 42°C. The membranes were

washed in 1×SSC, 0.1% SDS at 42°C. The moist membranes were exposed overnight on XAR-5 film (Kodak, Rochester, NY) at -70°C, using an intensifying screen.

Isolation of human ABP cDNA clones

Following the method of Benton and Davis (24), Eco RI fragments of the SHBG I cDNA clone were used to screen plaques from a human testis cDNA library cloned in lambda gt 11. Positive clones were plaque purified and the phage DNA extracted from plate lysates by conventional methods (20). The phage DNA was cut with Eco RI and Eco RI/Stu I, and the fragments were isolated by agarose gel electrophoresis (20). The restriction fragments of the inserts were cloned into M13 mp18/mp 19.

Cloning and analysis of the SHBG gene

A human genomic library constructed in a loric cosmid vector was the gift of P. F. R. Little (Chester Beatty Laboratories, London, UK). In screening the library, ³²P-labeled SHBG 1 cDNA was used as a probe. Recombinant loric cosmid DNA was isolated from liquid bacterial cultures and purified as previously described (25). The purified DNA was cut with several restriction enzymes and subjected to Southern blot analysis. A single Xba I fragment (4 kb), identified by hybridization with the ³²P-labeled SHBG 1 cDNA probe was isolated and sequenced by the shotgun method (26).

DNA sequence analysis

The dideoxy chain-termination method was used, with [α -³⁵S]thio-dATP (600Ci mmol⁻¹) and Sequenase[®], to sequence DNA cloned in M13 mp18/mp 19 (27, 28). Oligonucleotides were synthesized and used as primers to sequence regions for which the sequence had been determined on only one strand. Sequences were compiled and analyzed using computer software provided by R. Staden MRC Unit, Cambridge, UK (29, 30).

Primer extension analysis

5 µg poly(A)⁺ RNA was dissolved in 10 µl sterile water, and the RNA was denatured by the addition of 1 µl 100 mM MeHgOH. After 10 min, 2 µl each of 0.7 M β-mercaptoethanol (1.4 mmol) and RNasin (25 U) (Boehringer & Mannheim, Mannheim, West Germany) were added and allowed to react for 5 min. After addition of 1 µl ³²P-labelled oligonucleotide A or B (4×10⁶ cpm; 1.6×10⁷ cpm/pmol), 25 µl 1 M Tris-HCl, pH 8.3 (at 42°C), 3.5 µl 1 M KCl and 1 µl 0.25 M MgCl₂, the mixture was heated to 65°C during 2 min followed by a slow cooling to below 35°C. The primer extension reaction was commenced by the addition of 1 µl 25 mM dNTP:s and 1 µl AMV reverse transcriptase (10 U). The mixture was incubated at 42°C for 1 hour. The reaction was terminated by the addition of 1 µl 0.5 M EDTA. The reaction mixture was extracted once with phenol-chloroform (1:1) and ethanol precipitated. The precipitate was dissolved in 10 µl of 10mM Tris-HCl, pH 8.0, 0.1 mM EDTA. The reaction products were analyzed on 6 % polyacrylamide denaturing gels as used for DNA sequencing (27, 28). DNA sequencing reactions of characterized DNA fragments in M13 mp18 served as size markers.

RESULTS

Analysis of SHBG transcripts in the testis.

Northern blots of poly (A)⁺ RNA from human liver and testis were probed with a previously characterized SHBG cDNA derived from human liver (5). As shown in Figure 1, in both organs there seems to be a transcript with a size of approximately 1.6 kb. In order to further characterize the testicular transcript, a human testis cDNA library constructed in lambda gt 11 was screened with the SHBG cDNA. Out of 10⁶ plaques, nine hybridizing clones were isolated. By restriction mapping the positive clones could

Figure 1. Northern blot analysis of human SHBG mRNAs. mRNAs (20 μ g per lane) isolated from testis (T) and liver (L) were subjected to electrophoresis in denaturing formaldehyde-agarose gels and blotted onto a nylon membrane and hybridized with a 32 P-labeled liver SHBG cDNA. The mobilities of 32 P-labeled Hind III cleaved lambda DNA fragments are indicated.

Figure 2. Restriction map and sequencing strategy for testis SHBG cDNA and testis SHBGgrp cDNA. Arrows indicate the direction and extent of the fragment sequencing.

Pro	Arg	Phe	Lys	Gly	Ser	Pro	Ala	Val	Leu	Phe	Lys	Leu	Thr	Tyr	Ala	Val	Ile	Thr	Cys	20
CCG	CGG	TTC	AAA	GGC	TCC	CCC	GCA	GTG	CTT	TTT	AAA	TTG	ACA	TAT	GCA	GTG	ATA	ACC	TGC	60
B																				
Phe	Ser	Leu	Arg	Leu	Thr	His	Pro	Pro	Arg	Pro	Trp	Ser	Ala	His	Asp	Pro	Pro	Ala	Val	40
TTT	AGC	CTC	AGG	CTC	ACT	CAC	CCG	CCC	AGA	CCC	TGG	AGT	GCC	CAC	GAC	CCT	CCG	GCT	GTC	120
His	Leu	Ser	Asn	Gly	Pro	Gly	Gln	Glu	Pro	Ile	Ala	Val	Met	Thr	Phe	Asp	Leu	Thr	Lys	60
CAC	CTC	AGC	AAT	GGC	CCA	GGA	CAA	GAG	CCT	ATC	GCT	GTC	ATG	ACC	TTT	GAC	CTC	ACC	AAG	180
Ile	Thr	Lys	Thr	Ser	Ser	Ser	Phe	Glu	Val	Arg	Thr	Trp	Asp	Pro	Glu	Gly	Val	Ile	Phe	80
ATC	ACA	AAA	ACC	TCC	TCC	TCC	TTT	GAG	GTT	CGA	ACC	TGG	GAC	CCA	GAG	GGA	GTG	ATT	TTT	240
Tyr	Gly	Asp	Thr	Asn	Pro	Lys	Asp	Asp	Trp	Phe	Met	Leu	Gly	Leu	Arg	Asp	Gly	Arg	Pro	100
TAT	GGG	GAT	ACC	AAC	CCT	AAG	GAT	GAC	TGG	TTT	ATG	CTG	GGA	CTT	CGA	GAC	GGC	AGG	CCT	300
Glu	Ile	Gln	Leu	His	Asn	His	Trp	Ala	Gln	Leu	Thr	Val	Gly	Ala	Gly	Pro	Arg	Leu	Asp	120
GAG	ATC	CAA	CTG	CAC	AAT	CAC	TGG	GCC	CAG	CTT	ACG	GTG	GGT	GCT	GGA	CCA	CGG	CTG	GAT	360
Asp	Gly	Arg	Trp	His	Gln	Val	Glu	Val	Lys	Met	Glu	Gly	Asp	Ser	Val	Leu	Leu	Glu	Val	140
GAT	GGG	AGA	TGG	CAC	CAG	GTG	GAA	GTC	AAG	ATG	GAG	GGG	GAC	TCT	GTG	CTG	CTG	GAG	GTG	420
Asp	Gly	Glu	Glu	Val	Leu	Arg	Leu	Arg	Gln	Val	Ser	Gly	Pro	Leu	Thr	Ser	Lys	Arg	His	160
GAT	GGG	GAG	GAG	GTG	CTG	CGC	CTG	AGA	CAG	GTC	TCT	GGG	CCC	CTG	ACC	AGC	AAA	CGC	CAT	480
Pro	Ile	Met	Arg	Ile	Ala	Leu	Gly	Gly	Leu	Leu	Phe	Pro	Ala	Ser	Asn	Leu	Arg	Leu	Pro	180
CCC	ATC	ATG	AGG	ATT	CGC	CTT	GGG	GGG	CTG	CTC	TTC	CCC	GCT	TCC	AAC	CTT	CGG	TTG	CCG	540
Leu	Val	Pro	Ala	Leu	Asp	Gly	Cys	Leu	Arg	Arg	Asp	Ser	Trp	Leu	Asp	Lys	Gln	Ala	Glu	200
CTG	GTT	CCT	GCC	CTG	GAT	GGC	TGC	CTG	CGC	CGG	GAT	TCC	TGG	CTG	GAC	AAA	CAG	GCC	GAG	600
Ile	Ser	Ala	Ser	Ala	Pro	Thr	Ser	Leu	Arg	Ser	Cys	Asp	Val	Glu	Ser	Asn	Pro	Gly	Ile	220
ATC	TCA	GCA	TCT	GCC	CCC	ACT	AGC	CTC	AGA	AGC	TGT	GAT	GTA	GAA	TCA	AAT	CCC	GGG	ATA	660
Phe	Leu	Pro	Pro	Gly	Thr	Gln	Ala	Glu	Phe	Asn	Leu	Arg	Asp	Ile	Pro	Gln	Pro	His	Ala	240
TTT	CTC	CCT	CCA	GGG	ACT	CAG	GCA	GAA	TTC	AAT	CTC	CGA	GAC	ATT	CCC	CAG	CCT	CAT	GCA	720
Glu	Pro	Trp	Ala	Phe	Ser	Leu	Asp	Leu	Gly	Leu	Lys	Gln	Ala	Ala	Gly	Ser	Gly	His	Leu	260
GAG	CCC	TGG	GCC	TTC	TCT	TTG	GAC	CTG	GGA	CTC	AAG	CAG	GCA	GCA	GGC	TCA	GGC	CAC	CTC	780
*																				
Leu	Ala	Leu	Gly	Thr	Pro	Glu	Asn	Pro	Ser	Trp	Leu	Ser	Leu	His	Leu	Gln	Asp	Gln	Glu	280
CTT	GCT	CTT	GGG	ACA	CCA	GAG	AAC	CCA	TCT	TGG	CTC	AGT	CTC	CAC	CTC	CAA	GAT	CAA	GAG	840
Lys	Thr	Leu	Pro	Pro	Leu	Phe	Ala													288
AAG	ACT	CTT	CCA	CCT	CTT	TTT	GCC	TGA	ATGGCCTTTGGGCACAAGGTCAGAGGCTGGATGTGGACCAAGGC										910	
CCTGAACAGAAAGCCATGAGATCTGGACTCAGAGCTGCCCCAGAGGCCAGGCAATGGCACTGACGCTTCCCATTAAAGC																				989
TCCACCTAAGAACCCC																				

Figure 3. Nucleotide sequence of the human testis cDNA encoding the SHBG gene related protein (SHBGgrp), presented in the 5' to 3' direction. The deduced amino acid sequence is shown using the three-letter code. Numbers to the right of the sequences indicate amino acid residues (upper) and nucleotides (lower). The sequence differing from liver SHBG cDNA in the 5'- end is enclosed. The site of the frameshift generated by the splicing of exon VI to exon VIII is indicated by a star. The stop codon and a polyadenylation signal sequence are underlined as is the complementary sequence of oligonucleotide B (indicated by the letter B).

be separated into two groups, designated testis SHBG and testis SHBG-gene-related protein (testis SHBGgrp). The clone with the largest insert in each group was sequenced as outlined in Figure 2.

The sequence of the testis SHBG cDNA, which is 1112 bp long including a poly-A tail of 65 nucleotides (not shown), is identical to our previously described liver SHBG

Figure 4. Restriction map and sequencing strategy for the Xba I fragment of the cosmid clone SHBGg23 and for the Pst I fragment containing the SHBGgrp 5'-exon. Arrows indicate the direction and extent of the fragment sequencing. Filled boxes indicate exons and dotted lines unsequenced parts, the length of which were deduced by restriction mapping.

cDNA (5) with the following two exceptions: The testis cDNA is 41 bp shorter in its 5'-end and its poly-A tail starts four nucleotides closer to the stop codon than in the liver SHBG cDNA.

The sequence of the testis SHBGgrp cDNA is given in Figure 3. It is 1005 bp long and has an open reading frame extending from the beginning of the sequence to a stop codon (TGA) in positions 865–867 and thus coding for 288 amino acid residues. A comparison of the sequence with the longest published SHBG cDNA sequence, i.e. that of Hammond and coworkers (6), reveals the following: The first 96 nucleotides of the SHBGgrp sequence are not present anywhere in the SHBG cDNA, nucleotides 97–837 correspond to nucleotides 49–789 in the SHBG cDNA, and the remaining nucleotides (838–1005) correspond to nucleotides 998–1165 of the SHBG cDNA. In the only open reading frame, i.e. the one starting on the first nucleotide, nucleotides 97–835 of the SHBGgrp cDNA will be read in the same phase as in the SHBG cDNA but nucleotides 838–1005 will be read in a shifted phase giving rise to an entirely different and much shorter amino acid sequence than in the SHBG cDNA. There is no poly-A tail in the SHBGgrp cDNA but its sequence ends where the poly-A tail starts in our liver SHBG cDNA (5), which is 11 nucleotides earlier than in the cDNA of Hammond and coworkers (6).

Cloning and characterization of the human SHBG gene

We failed to isolate human SHBG gene clones from genomic libraries constructed in lambda phage (two Charon 4A libraries and one EMBL 3 library were screened of a total of 2×10^6 individual clones i.e., more than 2×10^{10} bp). Screening 50×10^3 recombinants ($> 2 \times 10^9$ bp) from a cosmid library yielded four hybridizing clones. The DNA was purified and subjected to restriction fragment analysis. Two of the clones showed identical patterns (SHBGg23 and SHBGg33), whereas the other two clones (SHBGg26 and SHBGg30) each had its own pattern. Southern blot analysis of Xba I digests of the clones SHBGg23, SHBGg30 and SHBGg33 all showed a single band of 4 kb hybridizing with both oligonucleotide A and oligonucleotide C, which are complementary to sequences in the 5'-end and the 3'-end of the SHBG cDNA, respectively. A fragment with the same size was seen when human leucocyte DNA was digested with Xba I and the Southern blot probed with the liver SHBG cDNA. Digestion of the fourth clone (SHBGg26) with Xba I yielded a 7 kb fragment that hybridized only to oligonucleotide C and it was concluded that SHBGg26 was a randclone.

The 4 kb Xba I fragment of SHBGg23 was isolated and sequenced according to the strategy outlined in Figure 4. From 70 randomly selected subclones in M13 mp 18, a contiguous sequence of 3810 nucleotides was obtained (Figure 5) with 96% of the sequence determined on both strands.

The coding part of the sequence consists of eight exons (exon I—exon VIII) separated by rather short introns devoid of any conspicuous features except for a sequence with similarity with the human Alu family of dispersed repeats (31) in intron six (nucleotides 2344 to 2651).

The first exon (exon I) codes for a typical signal peptide of 29 residues followed by 8 residues, which are the same as in the amino-terminal of the published protein sequence (4). The initiator methionine is encoded by nucleotides 361–363, and the polyadenylation signal consists of the nucleotides 3536–3541.

The sequences of the exons differ from the liver SHBG cDNA sequence of Hammond *et al.* (6) at positions 424–426 and 2951, of Que and Petra (7) at positions 2981, 3008, 3063, 3065, 3069, 3077, 3084, 3134, 3448, 3466 and 3550, and from all of the first 28 nucleotides of our cDNA sequence (5). This stretch of 28 enigmatic nucleotides can now be identified in the minus-strand of the gene (complementary to nucleotides 323–351 in the plus-strand, see Figure 5). The 5'-end sequence of our published cDNA (5) was thus an artifact, e.g. generated during the construction of the library. Nevertheless, it represents the most 5'-positioned sequence of a SHBG cDNA yet described.

Identification of genomic DNA encoding the 5'-end of testis SHBGgrp.

As can be seen in the gene sequence (Figure 5), the whole testis SHBGgrp transcript except for the first 96 nucleotides is derived from exons of the SHBG gene (exon II–VI, exon VIII). In order to find the genomic DNA (exon I') from which the first 96 nucleotides of the testis SHBGgrp cDNA is derived, oligonucleotide B, which is complementary to a sequence in the 5'-end of the testis SHBGgrp cDNA (Figure 3), was used to probe digests of the SHBGg23 clone. With Pst I a single hybridizing fragment of 0.6 kb was obtained. This fragment must have come from DNA localized 1.5 kb upstream of the translation start site in exon I, since a Pvu II digest of SHBGg23 yielded a 3.1 kb fragment hybridizing both with the above-mentioned oligonucleotide B and with oligonucleotide A, which is complementary to a sequence in exon I of the SHBG gene (see Figure 5).

The Pst I fragment was isolated and sequenced on both strands (Figure 4). The sequence (635 bp) is shown in Figure 6. The 96-nucleotide 5'-end of the testis SHBGgrp cDNA is recognized in positions 498–593. As no initiator methionine codon is present in the sequence of the Pst I fragment, at least one additional exon (a hypothetical exon 0) may exist further upstream in the genomic DNA.

Determination of sizes of the SHBG and SHBGgrp transcripts and their tissue distribution

Two oligonucleotides A and B, the first complementary to a sequence in the 5'-end of the SHBG mRNA and the other to a sequence in the 5'-end of the testis SHBGgrp mRNA, were used as primers in primer extension analysis with mRNA from liver and testis. Each oligonucleotide yielded a pattern of multiple extension products, which was identical with liver and testis RNA (Figure 7). The results suggest that transcripts for both SHBG and SHBGgrp are formed both in liver and testis, and that for each protein the transcript has a variation in start site but is identical in the two organs. The sizes of the transcripts, not counting the poly-A tails, were calculated to range between 1.43 and 1.55 kb for SHBG, and between 1.12 and 1.19 kb for SHBGgrp.

Potential mRNA start sites and promoter structure

The lengths in the 5'-direction of the primer extension products for SHBG (obtained with

oligonucleotide A) are indicated by markings above the genomic DNA sequence in Figure 5. Whether the positions of the markings represent real mRNA start sites is not clear. There are no conventional promoter elements such as TATA, CCAAT, CACCC or GGGCGG upstream of nucleotide 157, which is suggested by the markings to be the

```

      10      20      30      40      50      60      70      80      90     100
TCTAGACCTCAGGCCTGTGAATGCAAGGCTCCCCGAGTGACAGAAATCTTGGAGGACCTAGATCAGGCCCTAGAGGAGGAGAGGGGAGATGGAATATCC 100
      x
TCTCCAGATTACAGAACTTTCTCGGCAGTGGAGGATGATAGTGGAGGGGACTGTCTCTTACCCCATGATCCCCAGAGGGGTGATAGCTGAGTCTTGT 200
      x
GACTGGGCCCCGGGAGGGGTCAAGGGTCAGTGCCCTGTTTCTTTACCCCTCTCTCCCGGGCAACCTTTAAACCTCCACCGCCACACACAAGGCT 300
      Exon I
GCCTGCCTCTACACATTCTCCCAGAGTTGTCTGAGCCGCGGAGTGACAGTGGCTGATTATGGAGAGCAGAGGCCACTGGCTACCTCGCGCTGCTGC 400
MetGluSerArgGlyProLeuAlaThrSerArgLeuLeuL
euLeuLeuLeuLeuLeuLeuLeuArgHisThrArgGlnGlyTrpAlaLeuArgProValIleuProThrGln
TGTGTGCTGCTGTGCTACTACTGCGTCACACCCCGCAGGATGGGCCCTGAGACCTGTTCTCCCCACCCAGGTGCAGGAGCGGACAGGGCACTCAGCTC 500
      A
ATGCAGTCTTCCTCTCTCTCTCTGCGOCTGTAGCAGGGCTCTCCCTCTGTCTGTCTCTGACATGTCCTACTCAGCTTGTGTTGTTTCTCTTTCTGA 600
      Exon II
SerAlaHisAspProProAlaValHisLeuSerAsnGlyProGlyGlnGluProIleAlaValMetThrPheAspLeuThrLysIleThrLy
TAGAGTGCCCAAGACCTCCGGCTGTCCACCTCAGCAATGGCCAGGACAAGAGCCTATCGCTGTCTGACCTTTGACCTCACAAGATCACAAGTATG 700
GGGTGGCTAGCCCTGACCCAGTCCCTGGTCTGCGCTCTCTCCATCAGCTCTTCTTTTCCCTGCTCTCTTCTCTATCTGTGAACACCATCTC 800
      Exon III
CCCCAACCCACACTGGTTCTCAAAGGACATGACATACACAATCTTTCCTTCTGTGCTCTCCAGAACCTCTCTCTCTTGAAGTTCGAACCTGGGA 900
sThrSerSerPheGluValArgThrTrpAs
pProGluGlyValIlePheTyrGlyAspThrAsnProLysAspAspTrpPheMetLeuGlyLeuArgAspGlyArgProGluIleGlnLeuHisAsnHis 1000
CCGAGAGGGAGTGATTTTATGGGGATACCAACCTCAGGATGACTGGTTTATGCTGGGACTTCGAGAGCGCAGGCTGAGATCCAACCTGCAACATCAC
TrpAlaGlnLeuThrValGlyAlaGlyProArgLeuAspAspGlyArgTrpHisGln
TGGGCCAGCTTACGGTGGGTCTGGACACCGCTGGATGATGGGAGTGGCACCAGGTAAAGCTAGCTCTGTCTCAGGGAGGGATGTCTGGAGCTGG 1100
TCTGAGAAAGGGAAACAAAACCAAGTTATTGGCATCCCTTTAACCCTGTCTCTGTTTAAACACGAAACCCCAAGAGTAGCTATTCTTGGCCCC 1200
ATCTTTTCTGATGGGAATCTCAAGGCTCAGTCAGTATATAAGTGACAAGAGCTGAGTGACCAAGGCCAAGGATGCTAGCTGCTCTTTAAGGCATGTT 1300
      Exon IV
CTTTCACATATAGTACTAGGCTGCCTCAGGAAGGTGGCAGAAACAGATCCACGGGGCTCTGATTTGCTTCCACCTTCTCGAGGTGGAAGTCAAG 1400
ValGluValLys
MetGluGlyAspSerValLeuLeuGluValAspGlyGluGluValLeuArgLeuArgGlnValSerGlyProLeuThrSerLysArgHisProIleMetA 1500
ATGGAGGGGAGTCTGTGCTGCTGGAGTGGATGGGAGGAGTGTGCGCTGAGACAGTCTCTGGGCCCTGACCAACCAAGGCATCCCATCATGA
rgIleAlaLeuGlyGlyLeuLeuPheProAlaSerAsnLeuArgLeuPro
GGATTGGCTTTGGGGGCTGCTCTTCCCGCTTCCAACTCTGGTTGCGGGTAACTACACCCAGGGGTGGAACCTAGCCAAAGACTTGGTAAAGCACTG 1600
CTGGTGGCTGGCGTGGGAATCTAAGTCCACCTTTAGGAGAAAGGAGAGGTTGAGAGCTGCAAGGGGAGGCCAAATGCTCAGAGGGGAGTCAACT 1700
      Exon V
GAGGGCAGGGAGGTGCGGACTGCGCTCCGATGCGCTGATTCTACATCCCGTATCTTATCTCTGTACACTCCAGCTGGTTCTCGCCTGGATGGCTG 1800
LeuValProAlaLeuAspGlyCy
sLeuArgArgAspSerTrpLeuAspLysGlnAlaGluIleSerAlaSerAlaProThrSerLeuArgSerCysAspValGluSerAsnProGlyIlePhe 1900
CCTGGCGGGGATCTGTGCTGGACAAACAGGCGAGATCTCAGCATCTGCCCCACTAGCCTCAGAGCTGTGATGTAGAATCAATCCCGGATATTT
LeuProProGlyThrGlnAlaGluPheAsnLeuArgA
CTCCTCCAGGACTCAGGCAGAAATCAATCTCCGAGGTAGATTCTCGAGTCTATTTTCCACCCCTGGCCAGCTCAGCTGCCTCTGTCCCTCT 2000
      Exon VI
ACCATGGCCCTTTCTCTTCTGAGACCCAGCTTTGAGGCTCAGGATAATCATTCTCCACAGACATCCCGAGCTCATGCAGAGCCTGGGCT 2100
spIleProGlnProHisAlaGluProTrpAlaP
heSerLeuAspLeuGlyLeuLysGlnAlaAlaGlySerGlyHisLeuLeuAlaLeuGlyThrProGluAsnProSerTrpLeuSerLeuHisLeuGlnAs 2200
TCTCTTTGGACCTGGGCTCAAGCAGGCAGCAGGCTCAGGCCACCTCTGTCTTTGGGACACAGAGAACCATCTTGCTCAGTCTCCACTCCAAGA
pGln
TCAAGTAAAGGGGACAGTGGGGCATGCTGTATTCTAGTGGAGCTGGAGCAATGAGGAAGAGGAGTCCAACATGTCAATATTAGGAAGGTTTCCAGCC 2300
CAGGGAACATAACAGACTGGCTCCACAGAAATGTTTTTCAATTAATAATTAGCCAGGCATGGGCTGGTGTCTTGGCTGTAAATCCAGGTGCTGGAGGCA 2400
AGACCAGAGGATCAGCTTGAAGCCAGGAGTTTGACACAGCCTGGGCAACATAGCAGAGACCTCTGTCTAABAAAAAABAAAAATTAGCCAGGCATGGTGA 2500
CACATGCTCTGTGCGCTAGCTATTAGGAGCCTGAGGCAGGAGGTTCACTTGAGCCAGGAGTTTGAAGCTGCACTGAGCTATGATGTGCCACTGCACTC 2600
TGACCTGGGCCACAGTGAGCCCTGTCTCAAAAAATAAAAAATAAAAAAGGCTTATGATGGCACTCAGGTGGTGGTATGGGGCAGGGACATATCTT 2700
GAAGCTCCCAACAGCAAGCAACAGTTTTGACTTAGACTGCATATTACTTGGGCAAGTGTGGTTTCAAAAAGGGTCAGCCAAAAAATTTGGGGCAGG 2800
ATTTAAGTGGTGAAGTGGCAGTAGGTGGAGCATACGGAAGAGGCAATTAAAGCAGCTAGGGTGGAGGCCACAGCAGATAGGCCCGGCTCATTCTT 2900
      Exon VII
LysValValLeuSerSerGlySerGlyProGlyLeuAspLeuProLeuValLeuGlyLeuP
CCCTCTCTCTACCGTCCCTTCCACACACTCTCGAAGGTGGTGTGCTTCTGGGTGGGGCCAGGCTGGATCTGCGCTGGTCTTGGGACTCC 3000
roLeuGlnLeuLysLeuSerMetSerArgValValLeuSerGlnGlySerLysMetLysAlaLeuAlaLeuProProLeuGlyLeuAlaProLeuLeuAs 3100
CTCTCAGCTGAAGCTGAGTATGTCCAGGTGGTCTTGAGCCAAAGGTGAAGATGAAGGCCCTTGCCCTGCTCTTAGGCTGGCTCCCTCTCTTAA

```


nLeuTrpAlaLysProGlnGlyArgLeuPheLeuGlyAlaLeuProG
 CCTCTGGGCCAAGCCTCAAGGGCGTCTTCTCTGGGGCTTTACCAAGTAAAGAGAATGATGTTCAAGTTCATGAGCACAACATTGGAACAGCTCAAG 3200
 GGAGGCGGCACATTTTGGAGGGAAGAACTCTGGGAGGGAAGAATAAGGCCACAAGAAGATATGGGGGCGAGTGAAGGTAGTGCTTTTGCAAA 3300
 Exon VIII
 lyGluAspS
 CTCAGGTTGGAGGAGTGGAAAAGTGGGGAGAAGATTCTGGATCCGAGCCACCTTAATGCTCTAATGCCACCTTTGCACTACCTCCCTCTAGGAGAAGACT 3400
 erSerThrSerPheCysLeuAsnGlyLeuTrpAlaGlnGlyGlnArgLeuAspValAspGlnAlaLeuAsnArgSerHisGluIleTrpThrHisSerCy
 CTTCCACCTCTTTTGCCTGAATGGCCTTTGGGCACAAGGTACAGAGCTGGATGTGACCAGGCCCTGAACAGAAGCCATGAGATCTGGACTCACAGCTG 3500
 sProGlnSerProGlyAsnGlyThrAspAlaSerHis
 CCCCAGAGCCAGGCAATGGCACTGACGCTTCCCATTAAGCTCCAGCTTAAGAAGCCCCCTTTGAAAGTTACTGATTATTCATTTAATCAACAAATATT 3600
 C
 CACTGTGCACTAGCAATGTACAGGCACTGTGCGAAGTATTGAGTTGTCTTAATGAGCAAAAACACTCTGGTTCCTACCTCTTGGTGCCACGTCCTCAT 3700
 AGGGAAGCAGACATTCATCAAGGGCTAACTAATAAGTGGATAGTTGGAAGCACTGATAAGAAGAATTGGAGAGTTGTGAAAACATCTGGAGACTGGCCGG 3800
 GCGTGGTGGC

Figure 5. Nucleotide sequence of the human SHBG gene. The nucleotides are numbered from the Xba I restriction site and shown in the 5' to 3' direction. Each exon is numbered with a roman numeral and shown with its deduced amino acid sequence. The approximate positions corresponding to the lengths of the major primer extension products are marked (x) above the sequence. The sequence corresponding to the minus-strand of the previously described liver SHBG cDNA is indicated by a dotted line. The complementary sequences of oligonucleotides A and C are underlined and indicated by corresponding letters. In intron six a sequence, related to the human Alu family of dispersed repeats, is underlined.

transcription initiation site for the longest transcript (Figure 5). It cannot be excluded, therefore, that there is an additional untranslated exon further upstream, which may have promoter elements in its 5'-flanking region. If so, the 5'-boundary of exon I must be somewhere between nucleotide 157 and 360.

In Figure 6 the lengths in the 5'-direction of the primer extension products for SHBGgrp (obtained with oligonucleotide B) are indicated by markings above the genomic DNA sequence. As already suggested, additional exons may exist further upstream and the 5'-boundary of exon I' is unclear. The markings may have been placed in an intron and not correspond to real mRNA transcription start sites. If so, the 5'-boundary of exon I' should be somewhere between nucleotide 310 and 497, if judged from the length of the primer extension products, and most likely between nucleotides 401 and 402 according to the consensus for the splice acceptor sequence (32). In that case, the putative first exon of SHBGgrp (exon 0) should contain no more than 92 nucleotides.

DISCUSSION

Northern blot analysis of liver and testis mRNA clearly showed that both organs contain

10 20 30 40 50 60 70 80 90 100
 CTGCAGTCTTACCCGAATCAGCCTCAGGATATCTCCACAGTCTCCCTCCTTGGCCCTCTCGGATCCGACGGAAGCCATCCGGATCCCCGCTGTCTGGGA 100
 CCAAAAGTCCCGAGGGCTCGCAAAAGCGCACTAGACCCCTTAAAGGGCTACGGACTTGGATCCTGAAGAGCCTGAGAGAGCGGGTGGCGGGAGTCCGGG 200
 GGGAGCGCGGGTAGCGCGCCTGGTAAGTGGAGCTGGGATTCGGCGCGCTACGGGAGGAGAGATAGGCCAGCGAGGCGATCTCTGTCCGGCATAG 300
 x x
 CCCCACCCCTCGAATTCTCTCGCAGCAGGGGGCACAACCTGTGAGCAATCAGCTTGGAGAACAGGCAGCGCCGCTCCCCCAAGCCCCACCCCGACA 400
 Pro
 GCTGGATCTTGTGACTGGGCTCTGGGTAGAGTTCAAGGTTGGAGTGAAGCGGCTTCCTGCGGTTGTGTGGGTGTCCCAACCTGGGTGAGATACCCCG 500
 ArgPheLysGlySerProAlaValLeuPheLysLeuThrTyrAlaValIleThrCysPheSerLeuArgLeuThrHisProProArgProTrp
 CGGTTCAAAGGCTCCCCCGAGTCTTTTAAATTGACATATGACAGTGATAACCTGCTTTAGCCTCAGGCTCACTACCCGCCAGACCTGGGTAGGCC 600
 TTAAGACCTCAGCTCTGAAGCTGTTTCTCTGCAG

Figure 6. Nucleotide sequence of a Pst I genomic DNA fragment containing the alternative 5'-exon (exon I'). The sequence is shown in its 5' to 3' orientation. The 5'-end corresponds to a Pst I restriction site and the sequence ends with the 3' Pst I restriction site. The amino acid sequence is deduced from the testis SHBGgrp cDNA. The approximate positions corresponding to the lengths of the major primer extension products are marked (x) above the sequence.

Figure 7. Primer extension analysis of the 5'-end of human SHBG and SHBGgrp mRNAs. mRNAs from liver and testis were isolated and hybridized with ^{32}P -labelled oligonucleotide A or oligonucleotide B and primer extended with AMV reverse transcriptase. The products were subjected to denaturing polyacrylamide gel electrophoresis and autoradiographed. The size of the major products were determined by comparing with known fragment sizes of the genomic DNA. From left to right, the DNA ladders are T, C, G and A.

mRNA hybridizing with human SHBG cDNA, and that the mRNA is of the same size and is present in approximately equal amounts in the two organs. The size estimated from the northern blot (1.6 kb) is in acceptable agreement with the lengths of the mRNAs determined by primer extension analysis but somewhat smaller than the value 2 kb obtained for SHBG mRNA extracted from human liver and Hep G2 hepatoma cells and detected by probing with a ^{32}P -labeled oligonucleotide (35). Rat ABP mRNA has a size of 1.60 to 1.65 kb (17, 33, 34).

The demonstration of a correctly sized mRNA in testis hybridizing with liver SHBG cDNA, and the presence of a cDNA in a testis library with a sequence identical to that of liver SHBG cDNA, strongly suggests that the testis cDNA characterized here represents the protein commonly referred to as ABP.

Several previous studies have shown that the plasma protein SHBG and the ABP produced by Sertoli cells have many characteristics in common. The two human proteins have been shown to react with the same monospecific antiserum raised against human SHBG (16), and monoclonal antibodies raised against rat ABP crossreact with human

Figure 8. Schematic illustration of the organization of the human SHBG gene, its alternatively spliced transcripts and their respective proteins. The exons are numbered with roman numerals. The alternative 5'-exon is designated as I', and the putative exon containing a translation start codon is designated as 0. Non-identical regions in SHBG and the two feasible SHBG gene related proteins are shown by thin lines. In SHBG the proposed location of the steroid binding domain is indicated by S and the locations of the three carbohydrate groups are marked (●).

SHBG (36). The primary structure of rat ABP is 68% identical to human SHBG (18). Hybridization studies on cDNAs for rat ABP, human SHBG, and human ABP indicate that the two human cDNAs are more similar to each other than to the rat cDNA (33). Both human ABP and human SHBG are resolved into two bands, named the heavy and the light protomer (10, 16), when analyzed on SDS-PAGE.

There are also differences between SHBG and ABP, however. The sizes of the protomers differ slightly between the two proteins, and human plasma SHBG binds quantitatively to concanavalin A while human testis ABP exists in two forms, one binding to concanavalin A and the other not. SDS-PAGE reveals that both forms of ABP have subunits with two slightly different sizes which also SHBG has, but the sizes of the subunits of the two forms of ABP and their proportions are not exactly like those in SHBG. Furthermore, peptide mapping of human ABP and human SHBG has shown the form of ABP that does not bind to concanavalin A to differ from SHBG in some peptides, whereas the ABP that does bind to concanavalin A has a pattern that is more similar to that of SHBG (16). Two forms of ABP with similar properties have also been demonstrated both in the rat (37) and in the rabbit (37, 38). The differences observed between ABP and SHBG may be explained by differences in glycosylation, a hypothesis supported by deglycosylation experiments with rat ABP (39, 40).

Although the testis cDNA characterized in this work is not a full-length (it begins in exon II) and has a polyadenylation site different from that of the liver cDNA, it would appear that the protein backbones of the liver secreted SHBG and Sertoli cell secreted ABP are identical. If so, the chemical differences between the two proteins, referred to above, would be due to tissue-specific posttranslational modifications.

Comparison of the 5'-flanking sequences of the human SHBG/ABP and rat ABP mRNAs reveals a 67% identity in the first 36 nucleotides upstream from the start codon (ATG). This part corresponds to the non-translated leader sequence of rat ABP mRNA (41). Further upstream the similarity decreases to the random level. The region coding for the signal peptide in human SHBG consists of 87 nucleotides coding for 29 amino acids with a

sequence typical for signal peptides. There is a central core dominated by non-polar amino acids, including a sequence of 10 contiguous leucine residues. In the rat ABP signal peptide, which is 30 amino acid residues long, the corresponding part has 8 leucines.

In the screening of the testis cDNA library with the liver SHBG cDNA as probe we found a second cDNA that differs from the testis SHBG cDNA. A genomic DNA sequence coding for the 5'-end of this cDNA is located 1.5 kb upstream from the SHBG/ABP start codon. Part of the primary structure that may be deduced for the putative protein coded for by the alternative transcript is identical to that of human SHBG/ABP, and part of it is completely different (Figure 8). The putative protein, here called SHBG-gene-related protein (SHBGgrp) should have 247 amino acid residues in common with SHBG/ABP but differ otherwise (Figure 8). Depending on the length of the nontranslated leader sequence of the SHBGgrp transcript there should be a different amino-terminal of unknown length and the protein should be 109 residues shorter than SHBG/ABP in its carboxy-terminal region. Also theoretically, it should lack the two cysteine residues that have been proposed to form a disulfide loop in SHBG (4, 5). The protein probably has no steroid binding ability as it lacks the sequence encoded by exon VII, a part of the SHBG molecule that has been shown to participate in the steroid binding (2, 12, 42).

A striking analogy to the alternative splicing of the SHBG/ABP gene is found in the rat P-450 enzyme PB 1. Two forms of cDNA have been demonstrated for this enzyme (43). One contains 9 exons, the other 8; exon number 8 in the first form, which is known to carry part of the heme-binding domain, is deleted in the second form. Furthermore, as a result of the splicing of exon 7 to exon 9, there is a shift in the reading frame which produces an earlier appearing stop codon. The putative protein coded for by this nucleotide sequence is therefore truncated in its carboxy-terminal end as compared to the first enzyme protein; it thus lacks heme-binding capacity and does not function as a P-450 monooxygenase (43).

The result of the genomic Southern blot analysis makes the presence of more than one copy of the human SHBG/ABP gene improbable, which is in accordance with the single copy of the rat ABP gene (41).

The organization of the human SHBG gene is almost identical with that of the rat ABP gene (41). Both have the same number of exons and introns. With two exceptions, the introns of the two genes are of about the same length. Intron five is three times longer in the rat gene, and intron six in the human gene is twice the length of the corresponding rat intron. The exon/intron boundaries in the SHBG gene are located at exactly the same positions as in the rat ABP gene when the cDNA sequence is used as reference, and the sequences around the splice junctions are very similar in the two genes. No distinct TATA or CAAT boxes have been identified in the 5'-end of the human or the rat gene. Another shared feature is GA-rich and CT-rich clusters in the 5'-flanking region. But the somatostatin cAMP related regulatory element proposed to be present in the rat ABP gene (41) is not found in the human SHBG gene.

We have not been able to elucidate the regulation of the transcription of SHBG/ABP and SHBGgrp. If testis SHBGgrp is regulated by distinct promoter elements, one or more additional exons (exon 0) must exist further upstream from exon I' (Figure 8). Promoter elements may then be present 5' to this putative exon, and both the SHBG/ABP and the SHBGgrp transcripts may be regulated from this region. This would explain the lack of conventional promoter elements in the 5'-flanking region of the first SHBG/ABP exon. An alternative interpretation of the sequence data for SHBGgrp is that the alternative exon

I' is the first exon, which is a part of a non-translated leader sequence, and that the translation start codon is the first methionine codon in the second exon (exon II of the SHBG gene, nucleotides 667–669 in Figure 5). Since no distinct promoter element is found 5' to the longest transcript deduced from the primer extension experiments, the transcription may be promoted by other means analogous to those discussed for 'house-keeping' proteins (44).

ACKNOWLEDGEMENTS

This work was supported by grants from the Swedish Medical Research Council (Project No:s. B88-03X-5913 and B89-13X-08660-01A), the Faculty of Medicine at the University of Lund, the Albert Pålsson Foundation, the Greta and Johan Kock Foundation, the John and Augusta Persson Foundation, the Anna Lisa and Sven-Erik Lundgren Foundation, and the Cancer Research fund at Malmö General Hospital.

REFERENCES

1. Anderson, D.C. (1974) *Clin. Endocrinol.*, **3**, 69–96.
2. Petra, P.H., Namkung, P.C., Titani, K. and Walsh, K.A. (1986a) In Forest, M.G. and Pugeat, M. (eds). *Binding proteins of steroid hormones*. John Libbey, London, Paris.
3. Turner, E.E., Ross, J.B.A., Namkung, P.C. and Petra, P.H. (1984) *Biochemistry*, **23**, 492, 497.
4. Walsh, K.A., Titani, K., Takio, K., Kumar, S., Hayes, R. and Petra, P.H. (1986) *Biochemistry*, **25**, 7584–7590.
5. Gershagen, S., Fernlund, P. and Lundwall, Å. (1987a) *FEBS Lett.*, **220**, 129–135.
6. Hammond, G.L., Underhill, D.A., Smith, C.L. Goping, I.S., Harley, M.J., Musto, N.A., Cheng, C.Y. and Bardin, C.W. (1987) *FEBS Lett.*, **215**, 100–104.
7. Que, B.G. and Petra, P.H. (1987) *FEBS Lett.*, **219**, 405–409.
8. Van Baelen, J. Heyns, W. and DeMoor, P. (1969) *Ann. Endocr.*, **30**, 199.
9. Mischke, W., Weise, H.C., Graesslin, D., Rusch, R. and Tamm, J. (1979) *Acta Endocrinol.*, **90**, 737–742.
10. Cheng, C.Y., Musto, N.A., Gunsalus, G.L. and Bardin, C.W. (1983) *J. Steroid Biochem.*, **19**, 1379–1389.
11. Petra, P.H., Stanczyk, F.Z., Seneear, D.F., Namkung, P.C., Novy, M.J., Ross, J.B.A., Turner, E. and Brown, J.A. (1983) *J. Steroid Biochem.*, **19**, 699–706.
12. Hammond, G.L., Robinson, P.A., Sugino, H., Ward, D.N. and Finne, J. (1986) *J. Steroid Biochem.*, **24**, 815–824.
13. Gershagen, S., Henningsson, K. and Fernlund, P. (1987b) *J. Biol. Chem.*, **262**, 8430–8437.
14. Rosner, W., Aden, D. and Khan, M. (1984) *J. Clin. Endocrinol. Metab.*, **59**, 806–808.
15. Vigersky, R.H., Loriaux, D.L., Howards, S.S., Hodgen, G.B., Lipsett, M.B. and Chrambach, A. (1976) *J. Clin. Invest.*, **58**, 1061–1068.
16. Cheng, C.Y., Musto, N.A., Gunsalus, G.L., Frick, J. and Bardin, C.W. (1985) *J. Biol. Chem.*, **260**, 5631–5640.
17. Joseph, D.R., Hall, S.H. and French, F.S. (1986) In Forest, M.G. and Pugeat, M. (eds), *Binding proteins of steroid hormones*. John Libbey, London, Paris.
18. Petra, P.H., Titani, K., Walsh, K.A., Joseph, D.R., Hall, S.H. and French, F.S. (1986b) In Forest, M.G. and Pugeat, M. (eds). *Binding proteins of steroid hormones*. John Libbey, London, Paris.
19. Feinberg, D.P. and Vogelstein, B. (1983) *Anal. Biochem.*, **136**, 6–13.
20. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press. Cold Spring Harbor.
21. Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) *Biochemistry*, **18**, 5294–5299.
22. Aviv, H. and Leder, P. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 1408–1412.
23. Thomas, P.S. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5201–5205.
24. Benton, W.D. and Davis, R.W. (1977) *Science*, **196**, 180–182.
25. Little, P.F.R. and Cross, S.H. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3159–3163.
26. Bankier, A.T. and Barrell, B.G. (1983) In Flarell, R.A. (ed), *Techniques in Nucleic Acid Biochemistry*, vol. 85. Elsevier, Limerick, U.K.
27. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
28. Biggin, M.D., Gibson, T.T. and Hong, G.F. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3963–3965.
29. Staden, R. (1982a) *Nucleic Acids Res.*, **10**, 2951–2961.

30. Staden, R. (1982b) *Nucleic Acids Res.*, **10**, 4731–4751.
31. Jelinek, W.R. and Schmid, C.W. (1982) *Annu. Rev. Biochem.*, **51**, 813–844.
32. Breathnach, R. and Chambon, P. (1981) *Annu. Rev. Biochem.*, **50**, 349–383.
33. Reventos, J., Hammond, G.L., Crozat, A., Brooks, D.E., Gunsalus, G.L., Bardin, C.W. and Musto, N.A. (1988) *Mol. Endo.*, **2**, 125–132.
34. Reventos, J., Hammond, G.L., Bardin, C.W., Gunsalus, G.L., and Musto, N.A. (1986) In Forest, M.G. and Pugeat, M. (eds). *Binding proteins of steroid hormones*. John Libbey, London, Paris.
35. Mercier-Bodard, C., Radanyi, C., Roux, C., Groyer, M.T., Robel, P., Dadoune, J.P., Petra, P.H., Jolly, D.J. and Baulieu, E.E. (1987) *J. Steroid Biochem.*, **27**, 297–307.
36. Kovacs, W.J., Bell, B.W., Turney, M.K. and Danzo, B.J. (1988) *Endocrinology*, **122**, 2639–2647.
37. Hansson, V. (1981) *Int. J. Androl.*, **4**, 220–226.
38. Cheng, S.-L., Kotite, N. and Musto, N.A. (1984) *J. Steroid. Biochem.*, **21**, 669–676.
39. Danzo, B.J. and Bell, B.W. (1988) *J. Biol. Chem.*, **263**, 2402–2408.
40. Musto, N.A. (1988) *Ann. N.Y. Acad. Sci.*, **538**, 37–44.
41. Joseph, D.R., Hall, S.H., Conti, M. and French, F.S. (1988) *Mol. Endo.*, **2**, 3–13.
42. Petra, P.H., Que, B.G., Namkung, P.C., Ross, J.B.A., Charbonneau, H., Walsh, K.A., Griffin, P.R., Shabanowitz, J. and Hunt, D.F. (1988) *Ann. N.Y. Acad. Sci.*, **538**, 10–24.
43. Kimura, H., Yoshioka, H., Sogawa, K., Sakai, Y. and Fuji-Kuriyama, Y. (1988) *J. Biol. Chem.*, **263**, 701–707.
44. Dynan, W.S. (1986) *Trends Genet.*, **2**, 196–197.

<p>This article, submitted on disc, has been automatically converted into this typeset format by the publisher.</p>
--