MOLECULAR AND CELLULAR BIOLOGY, Oct. 1990, p. 5580–5585
0270-7306/90/105580-06$02.00/0
Copyright © 1990, American Society for Microbiology
Vol. 10, No. 10

The Promoter and First, Untranslated Exon of the Human Glucocorticoid Receptor Gene Are GC rich but Lack Consensus Glucocorticoid Receptor Element Sites

JIAN ZONG, JAVED ASHRAF, AND E. BRAD THOMPSON*

Department of Human Biological Chemistry and Genetics, The University of Texas Medical Branch, Galveston, Texas 77550

Received 6 March 1990/Accepted 10 July 1990

Glucocorticoid receptor mRNA is regulated by glucocorticoids. We found no consensus glucocorticoid response element, TATA box, or CAAT box but many GC boxes in ~3 kilobases of the 5′-flanking sequence of the human glucocorticoid receptor gene. We identified several transcription start sites, an untranslated exon 1, and the coding content of exon 2.

In a cell-specific manner, glucocorticoid receptor (GR) and GR mRNA levels may be up or down regulated by glucocorticoids (2, 6, 7, 10, 15, 18, 25, 30, 35, 38–42). The mechanisms of this regulation are unknown. To understand how steroid-regulated genes are controlled, it is necessary to know their primary sequences. Direct transcriptional up regulation by glucocorticoids seems to require one or more GR elements (GREs)—to which the ligand-activated GR binds—in the regulatory portion of the controlled gene (3, 4, 11, 22, 33, 44). To function, GREs require a 15-nucleotide sequence with partial dyad symmetry, e.g., GGTACANN NTGTTCT (11). The hexamer TGTTCT seems to be nearly invariant within this sequence. In 10 GREs found in six genes, this hexamer is found without alteration but for the single exception of an A substituted for a T at the 5′ position in one of the two GREs in tryptophan oxygenase (11). In the GGTACA hexamer, greater flexibility is allowed. However, the entire 15-nucleotide sequence appears to be required for function; half-GREs, consisting of one hexamer only, do not seem to mediate GR induction, probably because the GR acts as a dimer at its GRE site (3, 11). We therefore cloned and sequenced a considerable portion of the 5′-flanking region of the genomic human GR (hGR) gene. In doing so, we discovered and sequenced a previously unreported untranslated exon 1. By further sequencing, we found that exon 2 codes for the start of translation and the entire amino-terminal portion of the hGR, up to the DNA-binding region. We also identified multiple transcription start sites.

Genomic clones were obtained by screening a human leukocyte library in lambda-dash bacteriophage. About 10^6 phage were first screened with a 3-kilobase (kb) hGR cDNA coding region probe (13) labeled by nick translation (29). From recombinant phage that screened positive, two that contained an insert of about 20 kb and had the extreme 5′ end of the known coding sequence were identified by rescreening with a synthetic oligonucleotide whose antisense sequence corresponded to the untranslated hGR 5′ sequence +35 to +85, according to the published numbering (20). One genomic insert was cut into smaller fragments by use of restriction endonucleases, and the 1.6-kb *Sma*I fragment that still hybridized with the oligonucleotide probe was transferred to M13 phage and sequenced. A series of experiments allowed identification of several overlapping clones whose sequences comprised 2,996 bases of the 5′ flank of the gene for the hGR leading into exon 1. In addition, the continuous sequences through exon 1 and 439 nucleotides into intron 1 were determined (Fig. 1). M13 templates to be

```
-2996
AGAAAAAAAA ATGAAGGGAG AATAATTGGG AACTCTGTGA AATCATAAAT TATTTAAGCA GAGTACCTAC CACTAAGGAT TTGAGCATGT ATGAGCGATG
-2896
AGGTGGATTC TTAGAGCAGA GATTCAGCAG TGAGACAGGC TAATGGCATG GGGAGATGTG TTCTATTAGA CAACTGTATT GTCCCCTTAA TTATATACAT
-2796
ATATATGTAT ATAATATATA TGTGTATACA TATGTATACA TACATGTATA TATAAATAGA ATTCTTAATC ATTTTAAGCA GTGTACAAGG ATAATTATTA
-2696
GCATGGGCTA TTTAACTCAC TTTTAAAACG TGTAAAACAT TATTGTAGGG TTGGGTCTTG ATTTGTTTCC CATGAAACTG TTGGTAGCCC TGGGGCCAAA
-2596
TTAACGAAAG ACATCTCATT CTAGTGCTTG AGGCTCAGAA ATTGGAAAAA CGGGCAACTA GGTCAAGCAG TGTTCTCATA GCCTCACAGA TGAGCATCCA
-2496
AAGCAAGGGC CTCCTTGTAG TTGACTTGTG CCAAGGGACA GGAGAAGCGG AGTCTGCCTT CTGGGTCCAG AAGGGGGTTT TGTTAACATG GAGTTGCTCA
-2396
GCGCCCTTCA TAAAAATTCT TCTGGCTGAG GGTTCTACGT TGGCATACGG TTGGTTCCCT CTTCTTTTCC GAGGTGGCGA GTATCTCTTC CTTTGCCAAG
-2296
ATGGGCGCTC CAGAATCCTC TGGAGGCGGC CCCCGTAGAT CGTCTCCGGA CAAGAGGCTT GCTGAAAGCC TACTTCTTTC CTTTCACATC AGACAATGCA
-2196
CAGGGAACCT GTTACCCTTG AGAACCAAGG AAGGACGGCT TAGGCTACCC GCGATCGCAA CCTTTGCCAA GATGGTGGCC GCGGGGACGG GCTGGCGACA
-2096
CTGTACCCTA CCAAGATGGC GGCGGGCGGC TTCCGGGACG CGCTTCCCCA ATCGTCTTCA AGATGTCAGA GCAGGGGGAG CCGCCGTCAG TCTGAGCGCG
-1996
GGCGAGGTGA GAGAGTGGCT GTGGCGAGCG CCCGAGCAGG ATTAGGTGGA GCTGCGGCAG CCCCCGCCCG TGTCAGGAGC TGGCAACGGT GTCACCTGTG
-1896
GGGGCGCAAA AGTTACCTCC CCAAACCCTA AACCCACACA GCACAACCTT TCCCAGAGTC ACAAAAATCA TAATCTGTGC CGCACAAGGT AGGAGGCTCG
-1796
GTCCCGGCAT CGTCCAAGCC TTCCCGACGC GGCGAGCTGG GGAAGGGAGC TGGGGCGGGG GCTTCCCGAC GGGCACCCCT CGCCCCACGG CCCTCTCCTT
-1696
TCTCAGGACG GACCACGAGT TCCCTTCCCC TTGGACTGAG GGGGAAGCTC CTAACAGGAA CATCTGTAGG GAGTTGAACG CTGGCATTTT AAAGCTGCCT
-1596
GTATTTTGTT TTATTTGTAG GGGCAGGGGT CCTATGAACG TGATAGGGTG AGCAACGCAC AGAGTCGAGG GCAGCAAATG TCAAGATTCG GGGGTGGGGC
-1496
CTGCACCGGG AACTTGGACG CGGGCCCTGG CCGGGGTGGA AGAAGAGGTC AGGAGTTTCG GAAGGGGGGC TATATTTCGC CAGCAACTTA CTATTTCGCC
-1396
TGCAACTTGC TTTTAAGCCT GCCGCCCCCT GCTTTCCTTA ATCATAATAA TAAAAAAAAA GTGCAAAGAA ATCCAGCTCG CTGGAGGTTT TGCATTTGGC
-1296
GTGCAACTTC CTTCGAGTGT GAGCACATTG GGCGGGAGGG GTGGGGGTTG AACTTGGCAG GCGGCGCCTC CTTCTGCCGC CGCCGCCGCC TCGCAGACTC
-1196
GGGGAAGAGG GTGGGGGACG GTCGGGGCGC GGGGGAGGGT GGGTTCTGCT TTGCAACTTC TCTCCCAGTG CGAGAGCGCG GCGGCGGCAG CTGAAGACCC
-1096
GGCCGCCCAG ATGATGCGGT GGTGGGGGAC CTGCCGGCAC GCCGACTCCC CCCGGGCCCA AAGTACGTAT GCGCCGACCC CCGCTATCCC GTCCCTTCCC
-996
TGAAGCCTCC CCAGAGGGCG TGTCAGGCCG CCCGGCCCCG AGCGCGGCCG AGACGCTGCG GCACCGTTTC CGTGCAACCC CGTAGCCCCT TTCGAAGTGA
-896
CACACTTCAC GCAACTCGGC CCGGCGGCGG CGGCGCGGGC CACTCACGCA GCTCAGCCGC GGGAGGCGCC CCGGCTCTTG TGGCCCGCCC GCTGTCACCG
-796
CAGGGGCACT CGGCGCTTGC CGCCAAGGGG CAGAGCGAGC TCCCGAGTGG GTCTGGAGCC GCGGAGCTGG GCGGGGGCGG GAAGGAGGTA GCGAGAAAAG
-696
AAACTGGAGA AACTCGGTGG CCCTCTTAAC GCCGCCCCAG AGAGACCAGG TCGGCCCCCG CCCTGCCGCC GCCACCCTTT TTCCTGGGGA GTTGGGGGCG
-596
GGGGGCGAAG CGCGGCGCAC CGGGCGGGGC GGCCACGCCA GGGGACGCGG GCGTGCAGGC GCGTCGGGGC CGGGGTGGCG GGGCCCGCGC GGAGGGCGTG
-496
GGGGCAGGGA CCGCGGGCGC CCCTGCAGTT GCCAAGCGTC ACCAACAGGT TGCATCGTTC CCCGCGCCGC CGCGCGGCCC CTCGGGCGGG GAGCGGCCGG
-396
GGGTGGAGTG GGAGCGCGTG TGTGCGAGTG TGTGCCGCCG TGGCGCCGCC TCCACGCCGC TCCCCGCTCG GTCCCGCTCG CTCGCCCAGG CCGGGCTGCC
-296
CTTTCGTGTC CGCGCTCTCT TCCCTCCGCC GCCGCCTCCT CCATTTTGCG AGCTCGTGTC TGTGACGGGA GCCCGAGTCA CCGCCTGCCG TCGGGGACGG
-196
ATTCTGTGGG TGGAAGGAGA CCCGCACCGG AGCGGCCGAA GCAGCTGGGA CCGGGACGGG CACGCGCGCC CGGAAGCCCC GACCCGCGGA GCCCGGCGCG
-96                              ^                     ^                             ^      +1
GGGCGGAGGG CTGGCTTGTC AGCTGGGCAA TGGGAGACTT TCTTAAATAG GGCTCTCCCC CCACCCATGG AGAAAGGGGC GGCTGTTTAC TTCCTTTTTT
+5                                   ^                                                                 *
TAGAAAAAAA AAATATATTT CCCTCCTGCT CCTTCTGCGT TCACAAGCTA AGTTGTTTAT CTCGGCTGCG GCGGGAACTG CGGACGGTGG CGGGCGAGCG
+105        EXON INTRON
GCTCCTCTGC CAGAG|GTAAG AAGCGAGGCG GGAGGGGGCC GGGGCGCGCT CGCTCCCCCG AGGTGCCGCT GGGACCGGAG ACAACTCGGG GGCCGCCGCG
+205
GGAGCCTACA AACTTTTATT AGCCTCGGGG AGTGGGGGTG GGGGGCTGGC AAGGGCCGGG CGACGGTGAC GAAAGGGTAG CGCGCGGGTG ACAGCGCTGG
+305
CCTCTTCCTC TCCCTCCGCC GGCGTCCCTG GCCGGGCCGA GGGGGAGGAA CCTGACCTCG GACGGCGAGC GGAGCCCTGT CGAACTCCGG GGCTTCGAGC
+405
CTCTCATTCT CGCGGGAATC CTGGCCTCTT TTCTCCCCCT AGTGTCCCCT TTCCTCCAAG GGGGTCGCCC GACACCCGTT TTCGTGGTGA GACTAAGCCG
+505
CGTCTGAATT TTACTCGCCC GAATATTTCA CCCACCCGCC CCGAGCGCGA GCCC
```

FIG. 1. Nucleotide sequence of the promoter region, additional 5′-flanking region, exon 1, and beginning of intron 1 of the hGR. Underlines indicate consensus GC boxes, carets indicate transcription start sites determined by RNase protection and/or primer extension assays, +1 and * indicate the first base of cDNA found by Hollenberg et al. (20), and exon/intron indicates the junction between the end of exon 1 and intron 1.

* Corresponding author.

TABLE 1. Potential regulatory elements in the 5′-flanking region of the gene for the hGR[a]

Element	Consensus	Sequence in the gene for the hGR	Location (nucleotides)
TATA box	TATAAA	NF[b]	
CAAT box	CCAAT	NF	
GRE	GGTACAnnnTGTTCT	NF	
GRE(−)	ATYACNnnnTGATCW	NF	
Negative GRE	5′-CGTCCA	5′-CGTCCA	−1786 to −1781
		TGGACG	−1482 to −1477
Estrogen response element	AGGTCAnnnTGACCT	AGGCCGGGCTGCCCT	−309 to −295
AP-1-binding site	TGACTCA	TGACACA	−899 to −893
GC box	GGGCGG CCGCCC	GGGCGG	−2073 to −2068
		GGGCGG	−1744 to −1739
		GGGCGG	−1267 to −1262
		GGGCGG	−728 to −723
		GGGCGG	−722 to −717
		GGGCGG	−601 to −596
		GGGCGG	−575 to −570
		GGGCGG	−570 to −565
		GGGCGG	−413 to −408
		GGGCGG	−96 to −91
		GGGCGG	−20 to −15
		CCGCCC	−1933 to −1928
		CCGCCC	−1375 to −1370
		CCGCCC	−1094 to −1089
		CCGCCC	−969 to −964
		CCGCCC	−812 to −807
		CCGCCC	−665 to −660
		CCGCCC	−639 to −634
		CCGCCC	+541 to +546
Cyclic AMP response element	TGACGTCA	AGAGGTCA	−1453 to −1446
		TGCCGGCA	−1065 to −1058
		TGCCGCCA	−779 to −772
		TGACGGGA	−234 to −227
		TGCCGTCG	−211 to −204
NF-1-binding site	TGGnnnnnAGCCAAT	TCGAGGGCAGCAAAT	−1532 to −1518
		TGTCAGCTGGGCAAT	−80 to −66

[a] Consensus sequences used for these searches were obtained from the following sources: GRE, GRE(−), and estrogen response element, reference 3; AP-1, reference 27; GC box, references 5 and 9; cyclic AMP response element, references 1 and 32; NF-1-binding site, reference 23; negative GRE, reference 8. Sequences were analyzed by using the IBI DNA analysis program; 75% identity was chosen as the arbitrary cutoff for a site to be considered potentially present.
[b] NF, Not found.

sequenced were generated by the cyclone sequencing method (IBI, New Haven, Conn.), and sequencing was done as described by Sanger et al. (37). Because the organization of exons that precede and code for the amino-terminal end of steroid receptor family proteins varies somewhat between receptors and species (12, 17, 19, 21, 26, 28, 31, 34, 36, 43), all of exon 2, which contains the translation start site, and its intron-exon boundaries were sequenced as well (see Fig. 3). Comparison of the genomic sequence with the cDNA sequence showed that exon 2 was 1,196 nucleotides long, ending in a typical intron-exon junction sequence and followed by an extensive noncoding sequence. While internal exons are not usually this long, they sometimes are, and long coding exons 1 or 2 occur in some other members of the steroid receptor superfamily (12, 21, 28, 34, 36, 43).

Computer searches for several known *cis*-active regulatory elements were performed on the flanking DNA sequence (Table 1). In the flanking sequence obtained, the gene for the hGR lacks TATA and CAAT boxes and no consensus GRE was found. At −2526 and −2838, two perfect TGTTCT half-GREs were found. However, careful examination of the sequences on either side of them revealed no resemblance to the required dyad-symmetrical GGTA CA, spaced 3 nucleotides away. A computer search for GGTACA itself, with a 75% (≥5 of 6) identity requirement, found none in the 3-kb flanking sequence. Very recently, evidence was presented for a negative GRE sequence of 5′-CGTCCA at nucleotide −63 in the gene for pro-opimelanocortin (8). We found one 5′-CGTCCA sequence in the gene for the hGR at −1786 and the reverse antisense sequence 5′-TGGACT at −1482. More or less weak homologies for an estrogen response element, an AP-1 site, two NF-1 sites, and several cyclic AMP response elements were found. The most striking feature of the sequence is the quantity of GC-rich regions, which include 18 GC boxes or SP-1-binding sites (5, 9) in the 3 kb immediately 5′ to the

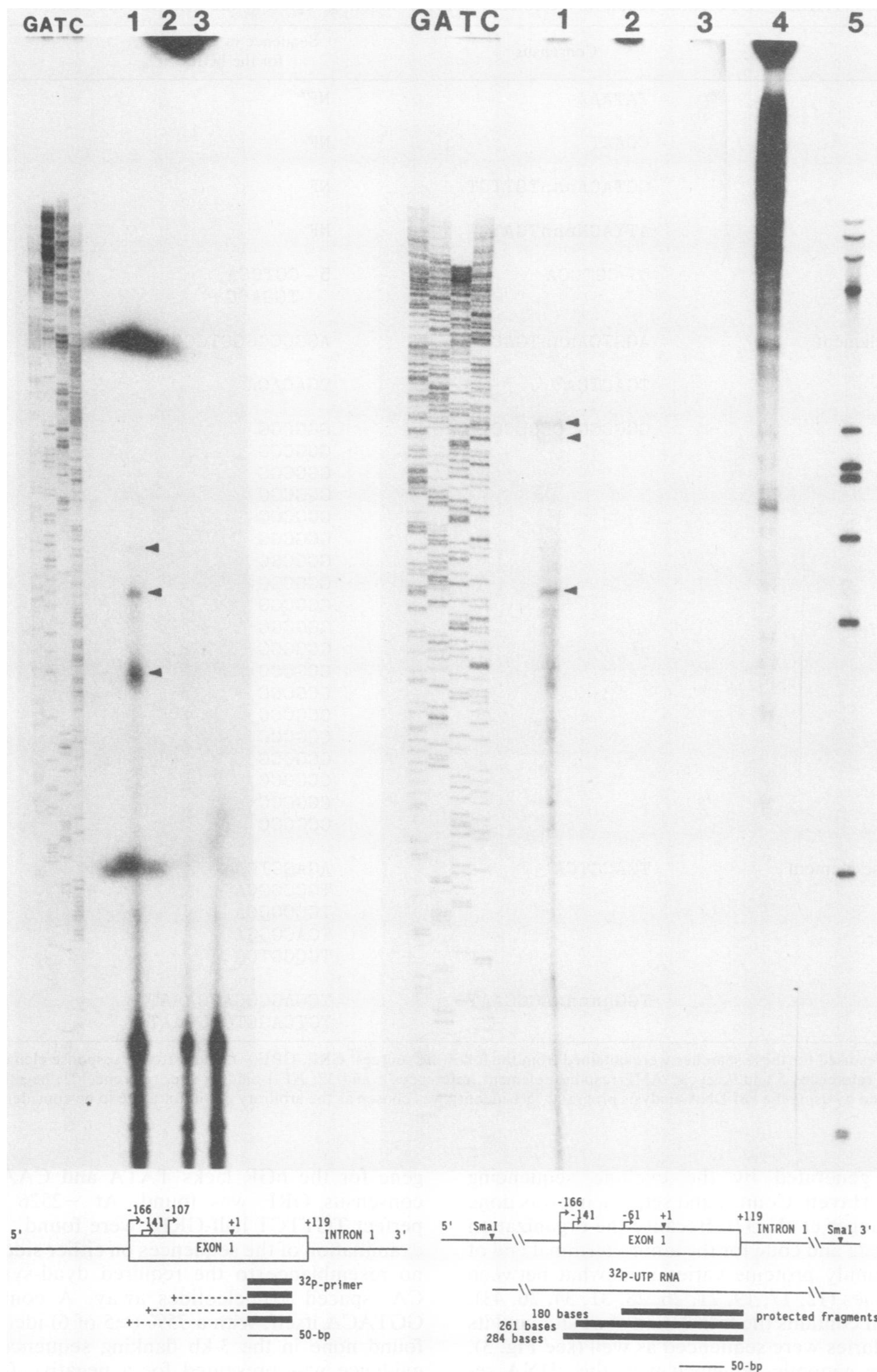

FIG. 2. Transcription initiation sites of the gene for the hGR. Primer extension analysis of human hGR mRNA is shown in the left panel. Lanes G, A, T, and C contained sequencing reaction products used as markers. Lanes: 1, extension products obtained by using total RNA prepared from Epstein-Barr virus-transformed lymphocytes; 2, the labeled primer itself; 3, extension products obtained by using control rabbit globin mRNA. A diagrammatic representation of these data is shown below the autoradiogram. RNase protection analysis is shown in the right panel. Lanes: 1, protected fragments obtained by using total RNA prepared from CEM C7 cells hybridized with a [^{32}P]UTP-labeled antisense RNA probe; 2, the labeled probe alone digested with RNases A and T1; 3, the labeled probe hybridized with control rabbit globin mRNA digested with RNase; 4, the labeled probe alone; 5, *Hae*III-digested lambda ϕX174 DNA markers. The arrowheads show the transcription initiation sites defined by either primer extension or RNase protection analysis. A diagrammatic representation of these data is shown below the autoradiogram. bp, Base pairs.

coding sequence. These are underlined in Fig. 1. Eight of the GC boxes lie in the reverse orientation, and eight, including two tandem repeats, are concentrated between nucleotides −560 and −820.

The transcription start sites of the gene for GR were determined by RNase protection (14) and primer extension (24) assays in two cell lines. For the former, a [^{32}P]UTP-labeled antisense RNA was made by cloning a 1.6-kb *Sma*I genomic fragment containing the promoter into plasmid pT7/T3-18 and transcribing it with T7 polymerase. Total cellular RNA from CEM C7 cells was hybridized with the antisense RNA probe and digested with RNases A and T1, and the products were analyzed by gel electrophoresis. To see whether the same start sites were used in other cells, we analyzed the mRNAs from Epstein-Barr virus-transformed human lymphocytes by using the primer extension method. The synthetic +35 to +85 oligonucleotide was end labeled with [^{32}P]ATP, hybridized with the cell RNA, and elongated with avian myeloblastosis virus reverse transcriptase, and the products were electrophoresed. Adjacent sequencing markers were used to read the transcription initiation sites. Several transcription start sites were seen by either method (Fig. 2). The four start sites seen are indicated in Fig. 1. The two sites representing the largest transcripts (at −141 and −166) were the same in both systems. Differing sites for initiation of a shorter transcript were also seen, at −61 in CEM C7 cells and at −107 in the Epstein-Barr virus-transformed cells. These may represent T (CEM C7)- versus B (Epstein-Barr virus-transformed)-cell differences in transcription due to cell-specific amounts or types of transcription factors. Multiple start sites are commonly seen in GC box-driven promoters.

A single clone that contained all of exon 2 and its intron-exon boundaries was sequenced (Fig. 3). This exon therefore contains the translation start site at the codon determined by bases +133 to −135. Exon 2 extends from nucleotides +120 to 1316 and consequently codes for the first 394 amino acids of the hGR, stopping short of the DNA-binding region (13). A similar long coding exon 1 has been described in the human androgen receptor gene (12).

Sequences of 5′-flanking regions of a few other steroid receptor genes have been published. The human estrogen receptor gene was found to have TATA and CAAT boxlike sequences at 27 and 103 nucleotides upstream from the major transcription start site (17). On the other hand, the promoter of the chicken progesterone receptor gene has been found to lack TATA and CAAT boxes and to contain several GC boxes (21). The preliminary account of research on the GR also suggested that its promoter was GC rich (16). NGFI-B and N-10 are two additional genes of the steroid receptor family whose ligands, if any, are unknown. NGFI-B has several GC boxes and no TATA box; N-10 contains a TATA-like sequence in its promoter (36, 43). The genes for NGFI-B, chicken ErbA, and rat Rev-ErbAα, like the gene for the hGR reported here, have untranslated first exons (26, 43, 45).

The GR itself is regulated by glucocorticoids, with its mRNA accumulating or diminishing in a cell-specific manner. Much current evidence indicates that for direct gene induction, glucocorticoids activate the GR, which binds to GREs in the responsive gene. Direct negative regulation may occur through GR binding to a GRE that interferes with other, more positively acting elements or to a negatively acting GRE (1, 8). Finding that the hGR contains no complete GRE sequence in an ~3-kb flanking sequence or exon 1 raises the possibility that the observed up regulation of GR

```
gaattctagc atagtattta tcaaatgtta tatgtattgg ttctcagaaa agcaaacagc cgtgattgaa aagaggtagg aattttaatg atcacacttc
cttttttga aattaaatac tttgacatca acttgaacct tcagaataat cagatgtaat gaattataat gtctgtgatt aacaaagcta cacgttcagt
gagcggcagg atgaatagcc aagcttagtt cgatacactt ttgccctcag ctgtgcaaat ggattgcatt gtacttttaa atctggcatg ctgaatggga
gcaggggaca tggcttttta ttctggaaga tagaaactac tcttctggta acaaagaatt tgattcggag ttaactaaaa ggttcattta acaagctgcc
                                                                                          120        130
tcttactaat cggatcagga agataatgtg actttagagc ttatgatgtt ttccccccgt ttttgttttt tgttttgtag TTGATATTCA CTGATGGACT
140        150        160        170        180        190        200        210        220        230
CCAAAGAATC ATTAACTCCT GGTAGAGAAG AAAACCCCAG CAGTGTGCTT GCTCAGGAGA GGGGAGATGT GATGGACTTC TATAAAACCC TAAGAGGAGG
240        250        260        270        280        290        300        310        320        330
AGCTACTGTG AAGGTTTCTG CGTCTTCACC CTCACTGGCT GTCGCTTCTC AATCAGACTC CAAGCAGCGA AGACTTTTGG TTGATTTTCC AAAAGGCTCA
340        350        360        370        380        390        400        410        420        430
GTAAGCAATG CGCAGCAGCC AGATCTGTCC AAAGCAGTTT CACTCTCAAT GGGACTGTAT ATGGGAGAGA CAGAAACAAA AGTGATGGGA AATGACCTGG
440        450        460        470        480        490        500        510        520        530
GATTCCCACA GCAGGGCCAA ATCAGCCTTT CCTCGGGGGA AACAGACTTA AAGCTTTTGG AAGAAAGCAT TGCAAACCTC AATAGGTCGA CCAGTGTTCC
540        550        560        570        580        590        600        610        620        630
AGAGAACCCC AAGAGTTCAG CATCCACTGC TGTGTCTGCT GCCCCCACAG AGAAGGAGTT TCCAAAAACT CACTCTGATG TATCTTCAGA ACAGCAACAT
640        650        660        670        680        690        700        710        720        730
TTGAAGGGCC AGACTGGCAC CAACGGTGGC AATGTGAAAT TGTATACCAC AGACCAAAGC ACCTTTGACA TTTTGCAGGA TTTGGAGTCT TCTTCTGGGT
740        750        760        770        780        790        800        810        820        830
CCCCAGGTAA AGAGACGAAT GAGAGTCCTT GGAGATCAGA CCTGTTGATA GATGAAAACT GTTTGCTTTC TCCTCTGGCG GGAGAAGCCG ATTCATTCCT
840        850        860        870        880        890        900        910        920        930
TTTGGAAGGA AACTCGAATG AGGACTGCAA GCCTCTCATT TTACCGGACA CTAAACCCAA AATTAAGGAT AATGGAGATC TGGTTTTGTC AAGCCCCAGT
940        950        960        970        980        990        1000       1010       1020       1030
AATGTAACAC TGCCCCAAGT GAAAACAGAA AAAGAAGATT TCATCGAACT CTGCACCCCT GGGGTAATTA AGCAAGAGAA ACTGGGCACA GTTTACTGTC
1040       1050       1060       1070       1080       1090       1100       1110       1120       1130
AGGCAAGCTT TCCTGGAGCA AATATAATTG GTAATAAAAT GTCTGCCATT TCTGTTCATG GTGTGAGTAC CTCTGGAGGA CAGATGTACC ACTATGACAT
1140       1150       1160       1170       1180       1190       1200       1210       1220       1230
GAATACAGCA TCCCTTTCTC AACAGCAGGA TCAGAAGCCT ATTTTTAATG TCATTCCACC AATTCCCGTT GGTTCCGAAA ATTGGAATAG GTGCGAAGGA
1240       1250       1260       1270       1280       1290       1300       1310
TCTGGAGATG ACAACTTGAC TTCTCTGGGG ACTCTGAACT TCCCTGGTCG AACAGTTTTT TCTAATGGCT ATTCAAGgta agatcagtgt ttttctgttt
cttaagaatg gtacatttaa ggtagattaa tagatgtaaa tcttcatgat ttatatgtgt tctctaaaga ttcatgtgct tttttatatg aataagttta
agtggccttt tgaaagtagg aaaggtagac aacctaagtg acatctgtcg taaccatttc aggtttttc cttaaatagt ggttttcagt atcccattgg
ccaacggtga ggattttatt taacattttt aaaataatgt tgctcattaa cagatatctt aacgaaaaat tatataaatt caggagagta taatgtctca
taatatcata ttgtgttgtg catggtcatt cagctgttta gaatatgttc ttatattaca ataaatgata ccctttctta catagtcaaa agttgtgctg
ccttatttgt aaattcgtta agtgttagct tgagattaaa gagttaaaag cagaagtact aacaaagagc cctattcttc aaactgaatc ttctgttaaa
gaatttgagt tttgaagttg ctaaagcaat gcagtgaaca gtgtaccaga ccatagtatt agacacaggt cttgctcaca gggttcttgc cataaagtag
acaagttatg tctgctgatg aatctcttta agagaggaat tggtgtcaac atggtgcaaa acaaaatttt acgttcaaat gttcctgcaa gttctcaagt
agataactga tggccaaaat tgttaagctt caattttcag ctttcgtttg atttttctct tttttttact cagtcgttta taagcatact gatatttttg
tctgacccaa aaaggtcaga aaatggaatt atcag
```

FIG. 3. Nucleotide sequence of exon 2 and an additional part of intron 1 and intron 2 of the hGR. Capital letters indicate exon sequences, and lowercase letters indicate intron sequences. Translation start site is at 133 to 135.

mRNA by glucocorticoids in certain cells (10, 15) is the result of secondary regulation. The two perfect half-GREs at −2526 and −2838 might act as binding sites for single GR molecules and thereby prevent the action of some positive factor, but they do not seem to do so by impinging on a cyclic AMP-responsive element, as described for the α subunit gene (1). Such half-GREs occur in other genes, up regulated by glucocorticoids, e.g., mouse mammary tumor virus, but do not act to mediate GR gene induction there (3, 11). It is generally felt that such free-standing hexamers are not positive regulatory elements for glucocorticoids. It is possible that the negative GRE at −1786 and the reversed negative GRE at −1482 could be utilized for direct control of the GR gene, and theoretically it is possible that functional GREs could be found even farther upstream or even in an intron. Either would represent a novel usage. However, the data presented here should enable workers to examine a key region of the gene for the hGR for regulatory function by direct tests.

Nucleotide sequence accession number. The GenBank accession number for the sequences shown in Fig. 1 and 3 is M32284.

This study was supported by Public Health Service grant 5R01 CA 41407 from the National Institutes of Health.

LITERATURE CITED

1. **Akerblom, I. E., E. P. Slater, M. Beato, J. D. Baxter, and P. I. Mellon.** 1988. Negative regulation by glucocorticoids through interference with a cAMP responsive enhancer. Science **241:** 350–353.

2. **Antakly, T., E. B. Thompson, and D. O'Donnell.** 1989. Demonstration of the intracellular localization and up-regulation of glucocorticoid receptor by *in situ* hybridization and immunocytochemistry. Cancer Res. **49:**2230s–2234s.
3. **Beato, M.** 1989. Gene regulation by steroid hormones. Cell **56:**335–44.
4. **Becker, P. B., B. Gloss, W. Schmid, U. Strähle, and G. Schütz.** 1986. *In vivo* protein-DNA interactions in a glucocorticoid response element require the presence of the hormone. Nature (London) **324:**686–688.
5. **Briggs, M. R., J. T. Kadonaga, S. P. Bell, and R. Tjian.** 1986. Purification and biochemical characterization of the promoter-specific transcription factor Sp1. Science **234:**47–52.
6. **Cidlowski, J. A., and N. B. Cidlowski.** 1981. Regulation of glucocorticoid receptors by glucocorticoids in cultured HeLa S3 cells. Endocrinology **109:**1975–1982.
7. **Danielsen, M., and M. R. Stallcup.** 1984. Down-regulation of glucocorticoid receptors in mouse lymphoma cell variants. Mol. Cell. Biol. **4:**449–453.
8. **Drouin, J., M. A. Trifiro, R. K. Plante, M. Nemer, P. Eriksson, and O. Wrange.** 1989. Glucocorticoid receptor binding to a specific DNA sequence is required for hormone-dependent repression of pro-opimelanocortin gene transcription. Mol. Cell. Biol. **9:**5305–5314.
9. **Dynan, W. S., and R. Tjian.** 1985. Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. Nature (London) **316:**774–778.
10. **Eisen, L. P., M. S. Elsasser, and J. M. Harmon.** 1988. Positive regulation of the glucocorticoid receptor in human T-cells sensitive to the cytolytic effects of glucocorticoids. J. Biol. Chem. **263:**12044–12048.
11. **Evans, R. M.** 1988. The steroid and thyroid hormone receptor superfamily. Science **240:**889–895.
12. **Faber, P. W., G. G. J. M. Kuiper, H. C. J. van Rooij, J. A. G. M. van der Korput, A. O. Brinkmann, and J. Trapmen.** 1989. The N-terminal domain of the human androgen receptor is encoded by one, large exon. Mol. Cell. Endocrinol. **61:**257–262.
13. **Giguere, V., J. M. Hollenberg, M. G. Rosenfeld, and R. M. Evans.** 1986. Functional domains of the human glucocorticoid receptor. Cell **46:**645–652.
14. **Gilman, M.** 1989. Ribonuclease protection assay, p. 4.7. *In* F. A. Ansubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology, vol. 1. John Wiley & Sons, Inc., New York.
15. **Gomi, M., K. Moriwaki, S. Katagiri, Y. Kurata, and E. B. Thompson.** 1990. Glucocorticoid effects on myeloma cells in culture: correlation of growth inhibition with induction of glucocorticoid receptor mRNA. Cancer Res. **50:**1873–1878.
16. **Govindan, M. V., M. Burelle, C. Cantin, M. Devic, C. Labrie, F. Labrie, Y. Lachance, G. Leblanc, C. Lefebvre, P. Patel, J. Simard, and U. Stropp.** 1987. Isolation and sequence of the human glucocorticoid receptor gene promoter. Cloning of the human androgen receptor cDNA. Serono Symp. Publ. Raven Press **1987:**187–209.
17. **Green, S., P. Walter, V. Kumar, A. Krust, J.-M. Bornert, P. Argos, and P. Chambon.** 1986. Human oestrogen receptor cDNA: sequence, expression and homology to v-*erb*-A. Nature **320:**134–139.
18. **Harmon, J. M., M. S. Elsasser, L. P. Eisen, L. A. Urda, J. Ashraf, and E. B. Thompson.** 1989. Glucocorticoid receptor expression in receptorless mutants isolated from the human leukemic cell line CEM-C7. Mol. Endocrinol. **3:**734–743.
19. **Hazel, T. G., D. Nathans, and L. F. Lau.** 1988. A gene inducible by serum growth factors encodes a member of the steroid and thyroid hormone receptor superfamily. Proc. Natl. Acad. Sci. USA **85:**8444–8448.
20. **Hollenberg, S. M., C. Weinberger, E. S. Ong, G. Cerelli, A. Oro, R. Lebo, E. B. Thompson, M. G. Rosenfeld, and R. M. Evans.** 1985. Primary structure and expression of a functional human glucocorticoid receptor cDNA. Nature (London) **318:**635–641.
21. **Huckaby, C. S., O. M. Conneely, W. G. Beattie, A. D. W. Dobson, M.-J. Tsai, and B. W. O'Malley.** 1987. Structure of the chromosomal chicken progesterone receptor gene. Proc. Natl. Acad. Sci. USA **84:**8380–8384.
22. **Jantzen, H. M., U. Strahle, B. Gloss, F. Stewart, W. Schmid, M. Boshart, R. Miksicek, and G. Schutz.** 1987. Cooperativity of glucocorticoid response elements located far upstream of the tyrosine aminotransferase gene. Cell **49:**29–38.
23. **Jones, K. A., J. T. Kadonaga, P. J. Rosenfeld, T. J. Kelly, and R. Tjian.** 1987. A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. Cell **48:**79–89.
24. **Kingston, R. E.** 1989. Primer extension assay, p. 4.8. *In* F. A. Ansubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology, vol. 1. John Wiley & Sons, Inc., New York.
25. **Lacroix, A., G. C. Bonnard, and M. E. Lippman.** 1984. Modulation of glucocorticoid receptors by mitogenic stimuli, glucocorticoids and retinoids in normal human cultured T cells. J. Steroid Biochem. **21:**73–80.
26. **Lazar, M. A., R. A. Hodin, D S. Darling, and W. W. Chin.** 1989. A novel member of the thyroid/steroid hormone receptor family is encoded by the opposite strand of the rat c-*erbA*α transcriptional unit. Mol. Cell. Biol. **9:**1128–1136.
27. **Lee, W., P. Mitchell, and R. Tjian.** 1987. Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. Cell **49:**741–752.
28. **Lubahn, D. B., T. R. Brown, J. A. Semental, H. N. Higgs, C. J. Migeon, E. M. Wilson, and F. S. French.** 1989. Sequence of the intron/exon junctions of the coding region of the human androgen receptor gene and identification of a point mutation in a family with complete androgen insensitivity. Proc. Natl. Acad. Sci. USA **86:**9534–9538.
29. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual, p. 109–110 and 320–328. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
30. **McIntyre, W. R., and H. H. Samuels.** 1985. Triamcinolone acetonide regulates glucocorticoid receptor levels by decreasing the half-life of the activated nuclear-receptor form. J. Biol. Chem. **260:**417–427.
31. **Milbrandt, J.** 1988. Nerve growth factor induces a gene homologous to the glucocorticoid receptor gene. Neuron **1:**183–188.
32. **Montimy, M. R., and L. M. Bilezikjian.** 1987. Binding of a nuclear protein to the cyclic-AMP response element of the somatostatin gene. Nature (London) **328:**175–178.
33. **Petersen, D. D., M. A. Magnuson, and D. K. Granner.** 1988. Location and characterization of two widely separated glucocorticoid response elements in the phosphoenolpyruvate carboxykinase gene. Mol. Cell. Biol. **8:**96–104.
34. **Ponglikitmongkol, M., S. Green, and P. Chambon.** 1988. Genomic organization of the human oestrogen receptor gene. EMBO J. **7:**3385–3388.
35. **Raaka, B. M., and H. H. Samuels.** 1983. The glucocorticoid receptor in GH_1 cells. J. Biol. Chem. **258:**417–425.
36. **Ryseck, R. P., H. Macdonald-Bravo, M. G. Mattei, S. Ruppert, and R. Bravo.** 1989. Structure, mapping and expression of a growth factor inducible gene encoding a putative nuclear hormonal binding receptor. EMBO J. **8:**3327–3335.
37. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
38. **Sapolsky, R. M., L. C. Krey, and B. S. McEwen.** 1984. Stress down-regulates corticosterone receptors in a site-specific manner in the brain. Endocrinology **114:**287–292.
39. **Schlecte, J. A., B. H. Ginsberg, and B. M. Sherman.** 1982. Regulation of the glucocorticoid receptor in human lymphocytes. J. Steroid Biochem. **16:**69–74.
40. **Svec, F.** 1985. Glucocorticoid receptor regulation. Life Sci. **36:**2359–2366.
41. **Svec, F., and M. Rudis.** 1981. Glucocorticoids regulate the glucocorticoid receptor in the AtT-20 cell. J. Biol. Chem. **256:**5984–5987.
42. **Tornello, S., E. Orti, N. F. DeNicola, T. C. Rainbow, and B. S. McEwen.** 1982. Regulation of glucocorticoid receptors in brain

by corticosterone treatment of adrenalectomized rats. Neuroendocrinology **35:**411–417.

43. **Watson, M. A., and J. Milbrandt.** 1989. The NGFI-B gene, a transcriptionally inducible member of the steroid receptor gene superfamily: genomic structure and expression in rat brain after seizure induction. Mol. Cell. Biol. **9:**4213–4219.
44. **Yamamoto, K. R.** 1985. Steroid receptor regulated transcription of specified genes and gene networks. Annu. Rev. Genet. **19:**205–215.
45. **Zahraoui, A., and G. Cuny.** 1987. Nucleotide sequence of the chicken proto-oncogene c-*erb*A corresponding to domain 1 of v-*erb*A. Eur. J. Biochem. **166:**63–69.