Pergamon

0960-0760(95)00115-8

J. Steroid Biochem. Molec. Biol. Vol. 53, No. 1–6, pp. 283–291, 1995
Elsevier Science Ltd. Printed in Great Britain

Structure and Regulation of the Luteinizing Hormone Receptor Gene

Maria L. Dufau,* Chon Hwa Tsai-Morris, Zhang Zhi Hu and Ellen Buczko

Section on Molecular Endocrinology, Endocrinology and Reproduction Research Branch, National Institute of Health, Bethesda, MD 20892, U.S.A.

Studies of the mechanisms controlling the expression of the rat luteinizing hormone receptor gene were pursued by characterization of the gene structure and identification of regulatory protein binding domains in the 5′-non-coding region of the gene and of 3′ non-coding functional domains responsible for generation of the major mRNA forms. The coding region of the rat LHR gene contains 10 introns and 11 exons, of which the first 10 exons comprise the hormone binding extracellular domain and exon 11, the seven transmembrane/G protein coupling module. Several alternative spliced variants of the LHR were identified that conform to deletions of complete and/or partial exons. Within the 6.2 kb of the 3′-non-coding region, two functional LHR pA domains (H1) and (H2) produce two sets of major mRNA transcripts, each coding for both holoreceptor and the form B splice variant. The H1 pA domain is unique to LHR and may represent a recombinant insertion domain. The functional efficiency of each pA domain is related to the specific pA signals, distal downstream elements, and tissue-specific factors. A TATA-less promoter region is present within the 173 bp 5′ flanking region of the gene, with Initiator (Inr) elements at transcriptional start sites. Transcription is dependent on the binding of the Sp1 protein at two Sp1 domains that each contribute equally to transcript initiation. Promoter activity is regulated by at least three additional DNA domains, R (−1266 to −1307 bp), C-box (−42 to −73 bp) and M1 (−24 to −42 bp) that bind multiple *trans*-factors in a tissue-specific manner. Basal promoter activity is enhanced by a functional M1 domain in LHR-expressing mouse Leydig tumor cells (MLTC) but not in non-expressing CHO cells. C-box binding factors either inhibit promoter activity or block inhibition through overlapping but not identical DNA binding domains that carry AP-2 and NF-1 elements. Removal of the AP-2 element within the C-box results in MLTC-specific transcriptional activation that may involve an MTLC M1/C-box interaction. In addition, competition for C-box factors by an upstream regulatory element (R) that is only inhibitory in CHO cells, indicates that both C-box binding factors compete for this upstream (R) domain in a tissue-specific manner. Competition between the inhibitory and neutral DNA binding factors within both upstream (R) and promoter domains (C-box) could provide a mechanism for the control of LH receptor gene expression in gonadal cells. These studies have revealed a complex pattern of transcriptional regulation that may reflect targets for signal-regulated changes in LH receptor gene expression.

J. Steroid Biochem. Molec. Biol., Vol. 53, No. 1–6, pp. 283–291, 1995

INTRODUCTION

The luteinizing hormone receptor (LHR) is a G-protein coupled glycoprotein located on the plasma membrane of gonadal cells. Binding to LH or hCG to the receptor triggers the activation of signal transduction pathways that regulate cell differentiation and steroidogenesis. The expression of LHR in the gonads is a hormonally regulated process. In the ovary, it is induced by FSH, estrogen and growth factors in the preovulatory follicle, is down-regulated following the ovulatory LH surge, and increases again during luteinization. In the testis, the LHR is expressed early in fetal development and remains fairly constant throughout adult life. This receptor can be down regulated by LH/hCG in adult but not in fetal LC [1, 2].

Proceedings of the IX International Congress on Hormonal Steroids, Dallas, Texas, U.S.A., 24–29 September 1994.
*Correspondence to M. L. Dufau.

Insights into the regulation and structure–function properties of the LHR have been greatly advanced by the availability of purified receptors [3], and most recently by the cloning of LHR cDNA [4, 5] and elucidation of its genomic structure [6, 7]. The LH receptor is a member of the seven transmembrane G-protein coupled receptor subfamily that includes the adrenergic and muscarinic receptors and many others. However, unlike several of these receptors, the LHR contains a large extracellular domain which displays the requisite conformation for high affinity binding to LH/hCG [8, 9]. This paper will review our recent work on the LH receptor gene structure and on the characterization of the promoter, upstream and regulatory elements, and the cognate transfactors involved in the regulation of the LHR gene transcription. It will also summarize our work on the 3′ UTR, including the identification of functional polyadenylation domains.

GENOMIC STRUCTURE OF THE LH RECEPTOR

The LHR gene spans at least 70 kb from the 5′ flanking region to the 3′-non-coding end and consists of 11 exons separated by 10 introns (Fig. 1). All introns are located in the extracellular domain. A single exon, exon 11, codes for the entire transmembrane domain, cytoplasmic tail and 3′untranslated region. The hormone binding domain encompasses most of the region spanning exons 1–10, and consists of two cys rich regions in exon 1 and 9 that border a leucine rich domain within exons 2–8 [6]. The general leucine motif of about 20 amino acids, repeated at approx. 70 bp intervals is separated genomically by introns. This region within the hormone binding domain may have originated from gene duplication and crossing over. The similarity in sequence, intron/exon structure and intron phase among the glycoprotein hormone receptors indicates that the extracellular binding domain arose from duplication of a common ancestral gene. Since exon 11 of the LHR and exon 10 of the FSH and TSH receptors contain the entire G-coupling seven transmembrane cytoplasmic domain, and is homologous in size and sequence to the intronless G-coupled receptors, the glycoprotein hormone receptor may have originated as a recombinant of the ligand binding extracellular domain and an ancestral intronless G-coupled receptor [8].

SPLICED VARIANT OF RECEPTOR FORMS, mRNA SPECIES, THE 3′UTR

The sequences of truncated forms of the LHR cDNA that have been identified in human [10], rat [6, 11] and pig libraries [4] are consistent with deletion of either complete or partial exons within the genomic structure. Truncated soluble LHR variants derived from rat, human and pig cDNA libraries result from splicing at various alternative acceptor splice sites in exon 11 with or without elimination of exon 9 and an ovarian human membrane variant derived from elimination of exon 9 was also reported [10]. A rat soluble truncated variant which results from splicing at an alternate acceptor site in exon 11 and lacks the transmembrane and cytoplasmic domains (form B) is expressed as a soluble secreted product with high affinity and specificity for LH and hCG [6, 8]. Based on mRNA abundance, this form is the major spliced form of the receptor in luteinized ovaries. In the ovary, among the three major mRNA species of 5.8, 2.6 and 2.3 kb, the 5.8 kb was predominant [12, 13]. These and a minor 4.4 kb species were present at most stages of ovarian maturation while three other minor species (8.0, 1.9 and 1.4 kb) were only observed at certain developmental stages [12, 13]. Two functional polyadenylation domains in the regions that are responsible for the major LHR mRNA species were identified at 2368–2491 bp (H1) for the 2.6 and 2.3 kb mRNAs (H1) and 5579–5768 bp (H2), for the 5.8 kb mRNA species (14) (Fig. 2). Multiple pA sites located 3′ of two identical copies of pA elements (AAUAUA in H1 and of AAUAAA in H2) accounted for microheterogenous poly(A) addition at each of the two pA regions. Of the two H1-terminated mRNA species, the 2.6 kb mRNA codes for the holoreceptor and the 2.3 kb for the splice variant form B with the observed differences consistent with 266 bp deletion in the spliced variant. The 5.8 kb ovarian mRNA species that are resolved in gels as a

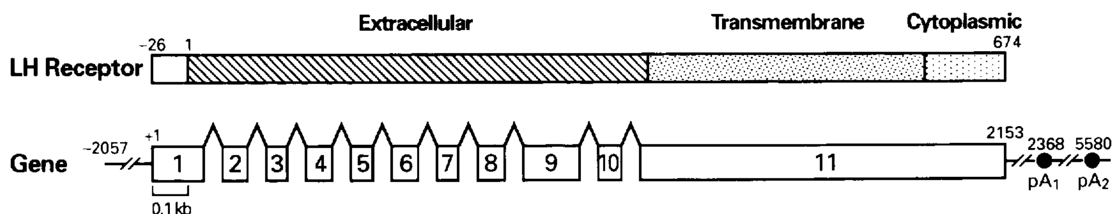

Fig. 1. Genomic organization of the rat LH/hCG receptor (LHR). Line 1, deduced amino acid sequence of extracellular, transmembrane, and cytoplasmic domains. Not drawn to scale. Amino acid (aa) positions begin with signal peptide (-26), +1 corresponds to Arg-27. Line 2, LH receptor gene including exons 1–11, introns ^, 2057 bp of 5′-flanking region, and 6120 bp of 3′-noncoding region, not drawn in scale [6, 8]. pA_1 and pA_2 correspond to functional polyadenylation domains H1 and H2 (see also Fig. 2).

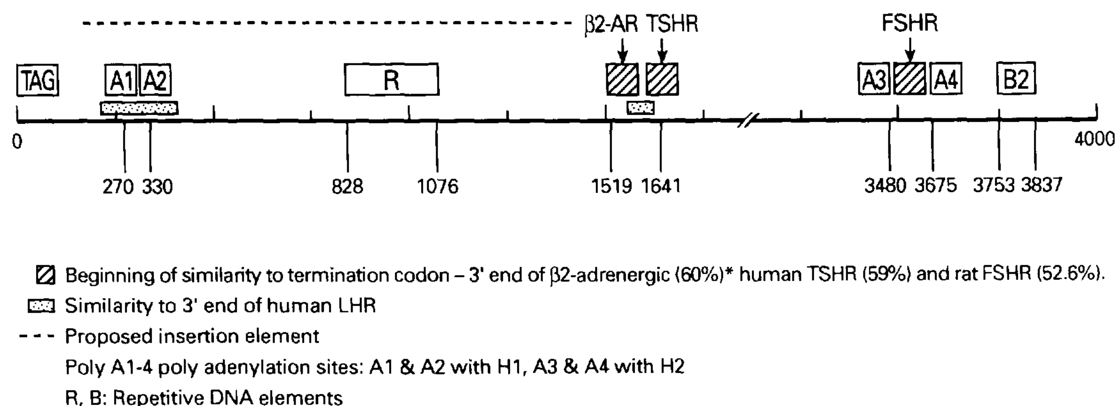

Fig. 2. Schematic representation of the organization of the 3′ non-coding region of the LHR gene. Regions of structural relevance derived from comparison of the rat LH receptor 3′ UTR sequences of the hLH, FSH, TSH, and β-adrenergic receptors R—a conserved repetitive rodent sequence not to be confused with the inhibitory R domain in the 5′ UTR of the LHR; B2: conserved rodent element [14].

broad mRNA band (5.2–6.2 kb), is derived from pA at the H2 domain and consists of transcripts for the full length as well as the truncated form B of the LHR. Direct sequence analysis demonstrates that the 5.8 kb and the 2.3/2.6 kb mRNAs contain about 3.5 kb and 0.3 kb of the 3′untranslated region, respectively [14]. The presence of spliced variant form B in both mRNA populations suggests that specific patterns of intronic splicing are independent of polyadenylation. A rodent repetitive DNA LINE R domain lies 3′ of H1 within the major 5.8 kb species and a B2 element was identified downstream of the H2 domain. The presence of the R domain within the 5.8 kb mRNA makes this form of the transcript susceptible to duplex formation with repetitive DNA and transcriptional regulation. The functional efficiency of each pA domain is related to the specific pA signals and is enhanced 5–7-fold in the presence of up to 2.6 kb of downstream genomic sequence 3′ to H1, or 1.3 kb 3′ to H2. Alignment of the 3′-non-coding region of LHR with TSH, FSH and β_2-adrenergic receptors indicates that the H1 pA signal and 3′ domain is unique to the LHR [14]. Thus, H1 may have arisen from a transposion in evolution (Fig. 2). The observed difference in steady-state levels between the H1 and H2 terminated transcripts may be of significance in regulation of LH receptor expression.

THE ORGANIZATION OF THE 5′ UTR

The LHR promoter, Sp1 functional domains

Serial deletion studies of the 5′-flanking region of the LHR gene have demonstrated that the promoter lies between −1 and −173 bp 5′ to the initiation colon. This promoter appears constitutively repressed by at least two domains located within sequences −173 to −626 and −626 to −990. A third repression region located between nucleotides −2056 and −1237 is observed only in non-expressing cells. The 173 bp promoter domain does not contain a TATA box, and may operate as an Sp1/Inr-directed transcriptional complex [6, 15]. In TATA-less genes, the mechanism of transcriptional initiation and its regulation is not uniformally established. In general GC-rich Sp1 DNA binding domains and initiator elements (Inr) have been proposed to act cooperatively to direct gene transcription [16]. Within the LHR promoter there are three Inr like elements ($Inr_{1,2,3}$) of the class found in the TdT gene (terminal deoxynucleotidyltransferase promoter) [17] which encompass the two major transcriptional start sites (TSS) at −14, −19 and a minor start site at −33 bp, within the Inr_3 element (Fig. 3). Of the four putative Sp1 consensus elements, only two functional Sp1 protein binding domains ($Sp1_2$ [−77 to −84] and

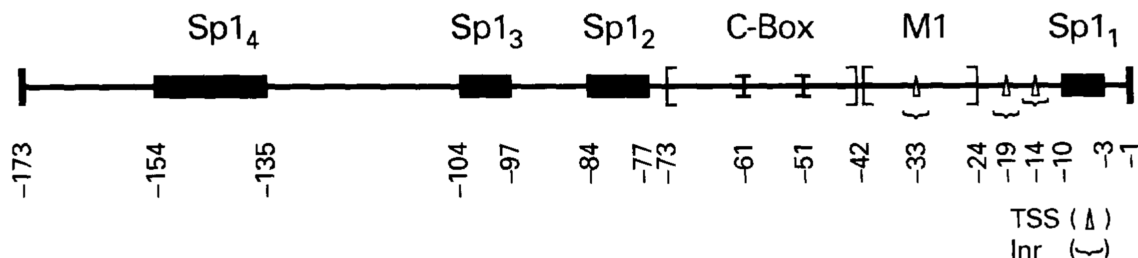

Fig. 3. Regulatory elements and regions within the 173 bp promoter domain of the LH receptor gene. TSS, transcriptional start sites; Inr, initiator elements.

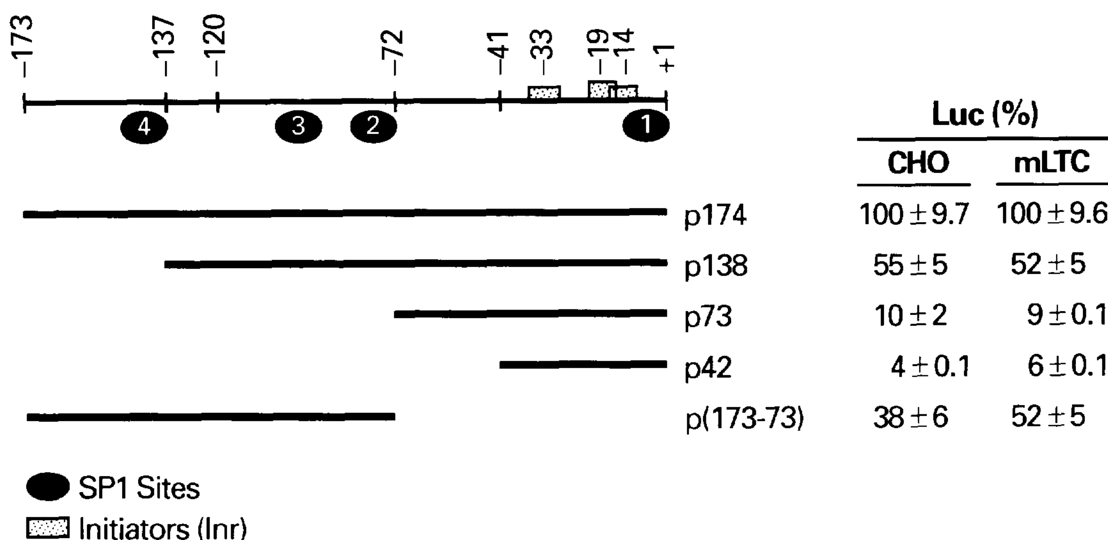

Fig. 4. Promoter region of the rat LHR gene. Luciferase activity of constructs from the LHR promoter. Constructs configured with designated fragments of the 173 bp promoter (+1 to −173 bp promoter) and subcloned in the pGL2-basic vector containing luciferase reporter gene were transfected in expressing (MLTC) and non-expressing CHO cells. Luciferase activity was determined in the cell extracts.

$Sp1_4$ [−135 to −154]), each contributing 50% of transcript activity, were identified by gel retardation analysis [15], deletion studies [18] and transfection studies in a Sp1 deficient cell line [19]. Deletion analysis of this 173 bp domain revealed that the 42 bp domain adjacent to the ATG intiation codon (which contains three transcriptional start sites) is not by itself sufficient to direct significant gene transcription. However, incremental addition of the domain about 30 bp upstream of the TSS sites containing the $Sp1_2$ binding domain induces significant gene transcription, with the upstream $Sp1_4$ binding domain required for maximal basal gene transcription. The 5′-flanking 137 nucleotide sequence that contains the protein binding $Sp1_2$ element, an activating M1 domain in the region −24 to −42 bp, and a regulatory C-box domain (−42 to −73 bp) was the minimal promoter domain capable of transcript initiation within this promoter region. The addition of sequences (−138 to −173 bp) that contain the $Sp1_4$ domain increased transcriptional activity by 100% (Fig. 4). Mutation of the $Sp1_1$ and $Sp1_3$ domains did not affect promoter activity, a finding which is consistent with gel retardation studies showing that these are not protein binding domains. The $Sp1_2$ and $Sp1_4$ domains are both of primary importance to transcript initiation, since basal promoter activity was reduced by 90% with simultaneous mutation of these Sp1 elements. $Sp1_4$-directed transcript initiation exhibits a level of 50% basal activity of the 173 bp promoter with mutation or deletion of the $Sp1_2$ domain in reporter gene constructs. Thus, activation by each of the $Sp1_2$ or $Sp1_4$ subpromoter domains reflects independent contributions by each element toward transcript initiation. Mutation of the Sp1 consensus sequence (GGGGCGGG) abolished $Sp1_2$ induced activity. However, this was not the case for the $Sp1_4$ domain, where mutation of the entire 19 bp, was found to be required to abolish $Sp1_4$ mediated transcriptional activity [18]. The Sp1 binding site in the $Sp1_4$ domain lies in the 5′G rich sequence [GGGGTGGGG] (subdomain **b**), rather than the 3′ GC box [GGGCGG] (subdomain **a**), and indicates a three trinucleotide element of GGG GTG GGG for the three Zn finger DNA binding Sp1 protein [19]. Mutation of nucleotides within the GC box indicates that non-Sp1 proteins are associated with this subdomain.

$Sp1_4$	GGGGTGGGGGGCGGGGAGA
Subdomain **b**	GGGGTGGGG
a	GGGGCGGGGAGA

The contribution of the GC box to LHR gene transcription appears to be tissue specific. Mutation of subdomain **a** or **b** in the 173 bp construct did not affect promoter activity in expressing cells, indicating that either subdomain can function independently to activate promoter activity as the wild type $Sp1_4$ domain. However, in non-expressing CHO cells, mutation of either subdomain reduced promoter activity by 50% indicating an additive effect between factors binding to these two domains in this cell type. The GC box (**a**) subdomain was demonstrated to be functionally associated with the downstream tissue specific M1 activator binding domain and functionally substituted the M1 domain in non-expressing cells but not in expressing cells. Cooperative effects between the Sp1 domain and other protein binding domains on the gene were demonstrated [19].

The M1 binding domain

The M1 binding domain which was identified as an activating domain in expressing MLTC, contains overlapping consensus elements for AP1-E, zestes, and the initiator element (Inr_3) from the terminal deoxynucleotidyltransferase gene family [15, 18]. In a reporter gene construct (p138GL) that contains $Sp1_2$, M1, C-box, and the native TSS but does not carry the $Sp1_4$ domain, the mutation of M1, designed to alter nucleotides common to consensus elements in this region, reduced transcriptional activity by 50% from the wild type in both CHO and MLTC cells. The subsequent mutation of an $Sp1_2$ site in the p138GL reporter gene construct further reduced transcription to near vector control levels in both cell types. This inhibition of reporter gene transcription caused by the M1 mutation was completely reversed in CHO cells, but not in mLTC, when the upstream (138–173 bp) domain containing $Sp1_4$ domain was added to the construct. This was dependent on a viable (**a**) subdomain within the $Sp1_4$ domain since mutation of this subdomain removed this activation only in non-expressing cells. Thus, the GC (**a**) subdomain is operative in the intact promoter only in CHO cells and in this cell type simultaneous mutation of both the GC box of $Sp1_4$ and the M1 subdomains is necessary to effectively inhibit $Sp1_4$-induced p174 promoter activity [18, 19]. The tissue specific differences in transcriptional activity exhibited by the M1/Sp1 mutant constructs may reflect differences in the nature of the M1 protein itself. Gel retardation analysis of the M1 protein–DNA complex with nuclear protein from MLTC and CHO cells shows that the properties of the M1 binding protein appear to be different in the two cell types with respect to migration rate and ion strength requirements of the DNA–M1 complexes. M1–DNA association is inhibited under low salt conditions in CHO cells, perhaps through a soluble binding factor that blocks the DNA binding domain of the *trans*-factor. However, the M1 *trans*-factor in the expressing MLTC cell functions as a constitutive activator, and its absence in CHO cells under low salt conditions may partially account for the absence of *in vivo* LHR gene transcription in this cell type [18].

Initiator elements

Deletion of the domain (+1 to −73) that contains the native transcriptional start sites of the LHR gene gave a major reduction in promoter activity of 50%, although both $Sp1_2$ and $Sp1_4$ elements were still present in the construct (Fig. 4). Thus, elements surrounding the transcriptional start sites in the 74 bp domain are important but not essential to the initiation complex/polymerase activity [18].

Mutation of CTCANTCT Inr sequences of the two major TSS (−14 and −19 bp) resulted in no change in transcriptional activity, and primer extension revealed that initiation takes place from the mutated element at the −14 and −19 bp positions. These results confirm that the Inr domains are not protein binding elements in the LHR gene although they may function in the polymerase–initiation complex.. This does not preclude that a different consensus Inr domain functions in the LHR gene, although a certain flexibility in the TSS is indicated by mutagenesis and deletion of the 74 domain and has been noted in previous studies with the terminal deoxynucleotidyltransferase Inr element [15]. In contrast, mutation of the third minor TSS gave a 50% inhibition in transcriptional activity, and this was not enhanced with the combination of mutants. However, I_3 falls within the tissue-specific M1 binding domain that also carries consensus elements for the AP-1 and zestes transfactors, so that the I_3 contribution cannot be separated from the M1 contribution. Since mutation of only the minor TSS showed reductions in reporter gene transcription it seems that this may not be due to defects in transcript initiation or the Inr element [18].

The C-box domain

The 30-nucleotide protein binding C-box element located (−43 to −73 bp) downstream of the $Sp1_2$ domain was divided into three segments (C1, C2, and C3) for characterization of functional studies. The C2 domain contains an AP-2 consensus element and an element similar to the NF-1 consensus sequence which spans the C1 and C2 domains (half palindrome in each region) and overlaps the AP-2 element.

C-Box Binding Domain

		←C1→	←C2→	←C3→	
LHR	−73	AGGGTGCTGGCAG	CCCCCAGGCG	GTCCAGCA −43	
AP-2 (Std)		ACGG	GCCGCGGGCG	GTCAGTTC	GATC
NF-1 Std)		CCTTTGGCAT	GCTGCCAATA	TG	

Gel retardation with labeled AP-2 consensus oligonucleotides reveals that *trans*-factors in CHO and MLTC nuclear extracts bind to the consensus AP-2 element and that these proteins are competed by oligonucleotides that contain the LHR AP-2 element in the C2 domain. Similarly *trans*-factors that bind to the consensus NF-1 element are competed by oligonucleotides that contain the LHR NF-1 element in the

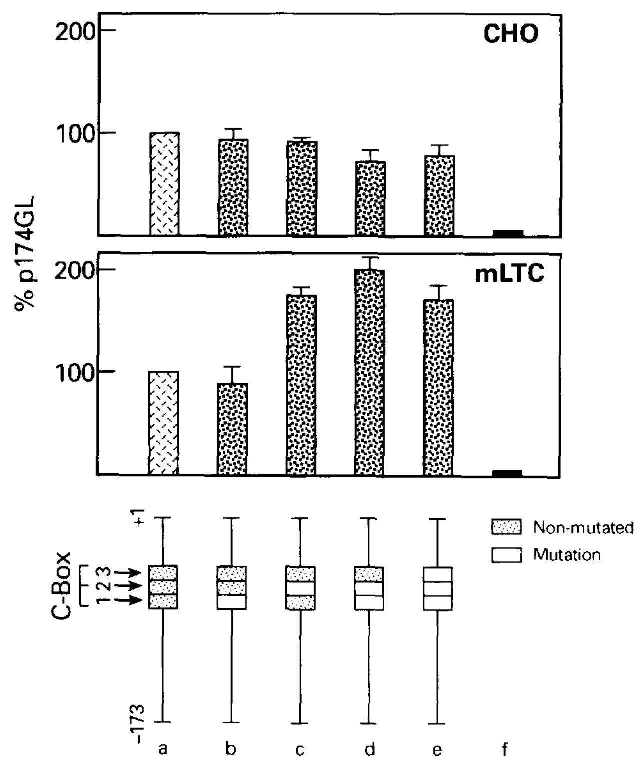

Fig. 5. Effect of mutations of the C-box domain on promoter activity of the LH receptor gene. Relative luciferase activity of wild type and mutant p174GL construct (+1 to −173) transfected in MLTC and CHO cells. Mutagenesis of the AP-2 element in C2 of the C-box resulted in significant (1-fold) increases in transcriptional activity when compared with the wild type construct. The C-2 DNA binding protein (AP-2) within the 173 bp promoter appears to be a repressor only in the expressing cell [18].

C-box with mutated AP-2 elements. Addition of polyclonal AP-2 anti-rabbit antisera to MLTC or CHO nuclear protein resulted in supershifts of DNA–protein complexes formed by labeled C-box or AP2 elements. AP2 binding is maximal to the C2 component of the C-box while the C1 and C3 components of the C-box appear to significantly decrease human AP-2 binding to the C-box. This indicates a potential for dissociation of a *trans*-factor binding from the AP-2 element that is contributed by either the C1 or C3 domains [18].

The C2/AP-2 DNA interactions can be associated with repression in MLTC, since specific mutations of the AP-2 element in C2, and presumably loss of AP-2 binding, resulted in promoter activation. Mutagenesis of the AP-2 element in C2 of the C-box resulted in significant (100%) increases in the transcriptional activity in MLTC when compared with the wild type construct, and the C2 DNA binding protein (AP-2) appears to be a repressor protein in this cell type. No further changes were observed with additional mutation of C1 or C3 in MLTC (Fig. 5). In contrast, in CHO cells, mutation of the C2 domain did not significantly change transcriptional activity and thus the CHO *trans*-factors that retard migration of C-box DNA appear to be neutral in function [18].

The MLTC-specific transcriptional activation also requires a functional M1 domain located 3′ adjacent to the C-box, and we considered that MLTC-specific activation might involve M1/C-box *trans–trans* interactions. Gel retardation/competition studies show a competition between M1 and C-box DNA for common MLTC nuclear protein(s)–DNA complexes but not for CHO nuclear protein–M1 DNA complex(es) [18]. The MLTC-specific M1/C-box basal promoter interaction appears to involve *trans*-factor binding to the LHR C1 domain rather than the LHR C2 AP-2 element. An 11-nucleotide element C1 was necessary to induce competition for the M1–DNA complexes. The 5-nucleotide element in C1 that contains the NF-1 half element only partially competed for M1 element, indicating that a putative C1 *trans*-factor that interacts with the M1 protein does not solely bind to the NF-1 half-element. The functional activation in the MLTC mutant that lacks the C2 AP-2 *trans*-factor may be envisioned as an M1/C1 *trans–trans* interaction that is sterically favored by the elimination of the AP-2 site and AP-2 binding protein in between M1 and C1. Thus, a general mode of transcript activation in MLTC may involve the initial release of a C2 repressor through an as yet unidentified *trans*-factor leaving the activated 5′ Spl_2 initiation complex that may include the C1–M1 protein complex. Steric interference by 3′ DNA binding protein factors is a common mechanism used in the inhibition of polymerase activity [18].

These and other findings suggest a differential influence of each subdomain on non-expressing vs expressing promoter activity. In MLTC, the C-box binding factors either inhibit promoter activity by binding to the AP-2 site or block inhibition through the overlapping NF-1 site. In CHO, C-box binding factors display a neutral effect on promoter activity. A related upstream sequence appears to control promoter activity in a similar manner in the non-expressing cells.

Upstream regulatory R domain

The basal 173 bp promoter activity of LHR gene is repressed by upstream inhibitory domains between 173 and 2057 bp. Deletion of sequences between −173 to −1237 reversed inhibition in MLTC cells, but failed to reverse inhibition in CHO cells. An upstream inhibitory domain that is exclusive for the non-expressing cells has been localized to a 40 bp sequence termed R domain at −1207 to −1266 bp. Ligation of the R domain to the 173 bp promoter caused inhibition of expression in non-expressing cells (CHO cells) but not in the expressing cell. The relevant protein binding domain in this sequence was localized to the 20 bp R2 subdomain (−1266 to −1275) by competitive gel retardation analysis. The R2 subdomain has similar

elements to those found in the C-box; an AP-2 like element and an overlapping palindromic NF-1 like element (Fig. 6).

Competition gel retardation analysis revealed that the transfactor binding to the R domain was specifically competed by the C-box sequence in the presence of CHO and MLTC nuclear proteins. The NF-1 like factor binds to this upstream R domain and is responsible for the lack of inhibition by this upstream sequence in expressing cells. In contrast, binding of AP-2 site to this upstream R domain is responsible for inhibition of promoter activity of the LHR gene in non-expressing cells. When an NF-1 like neutral factor, rather than an inhibitory AP-2, binds to the upstream R domain, regulation of transcriptional activity then occurs at the level of the promoter in expressing cells. In non-expressing CHO cells, both C-box *trans*-factors AP-2 or NF-1 display no regulation of promoter activity in the 173 bp region. The tissue-specific inhibitory effect of the upstream R domain on promoter activity in non-expressing cells is mediated exclusively through the binding of an inhibitory AP-2 factor (Fig. 7).

A working model of basal LHR transcript initiation based on the present findings can be represented by an Sp1-driven polymerase activity that is modulated by upstream *trans*-factors, including the R domain, and downstream *trans*-factors on the C-box/M1 domain (Fig. 8). These factors appear to be tissue-specific and either inhibit promoter activity or block inhibition through overlapping but not identical DNA binding domains. The overlapping protein binding DNA domains on the C-box (AP-2 and NF-1) AP-2 factors that can potentially be modulated by cAMP and phorbol esters have been observed to bind two regions of the gene (the upstream R and promoter-based C2 domains). Competition studies show that the neutral C1 factor (NF-1) can also bind to the upstream R domain in MLTC and that the inhibitory C2 factor (AP-2) can bind to the R domain in CHO cells (Fig. 9). There is no evidence of complex formation between different *trans*-factors on the R and C domains since supershifts were not observed with gel retardation. However, it is feasible that a putative initiation complex can bind the same protein component residing at either the promoter or upstream domain, with

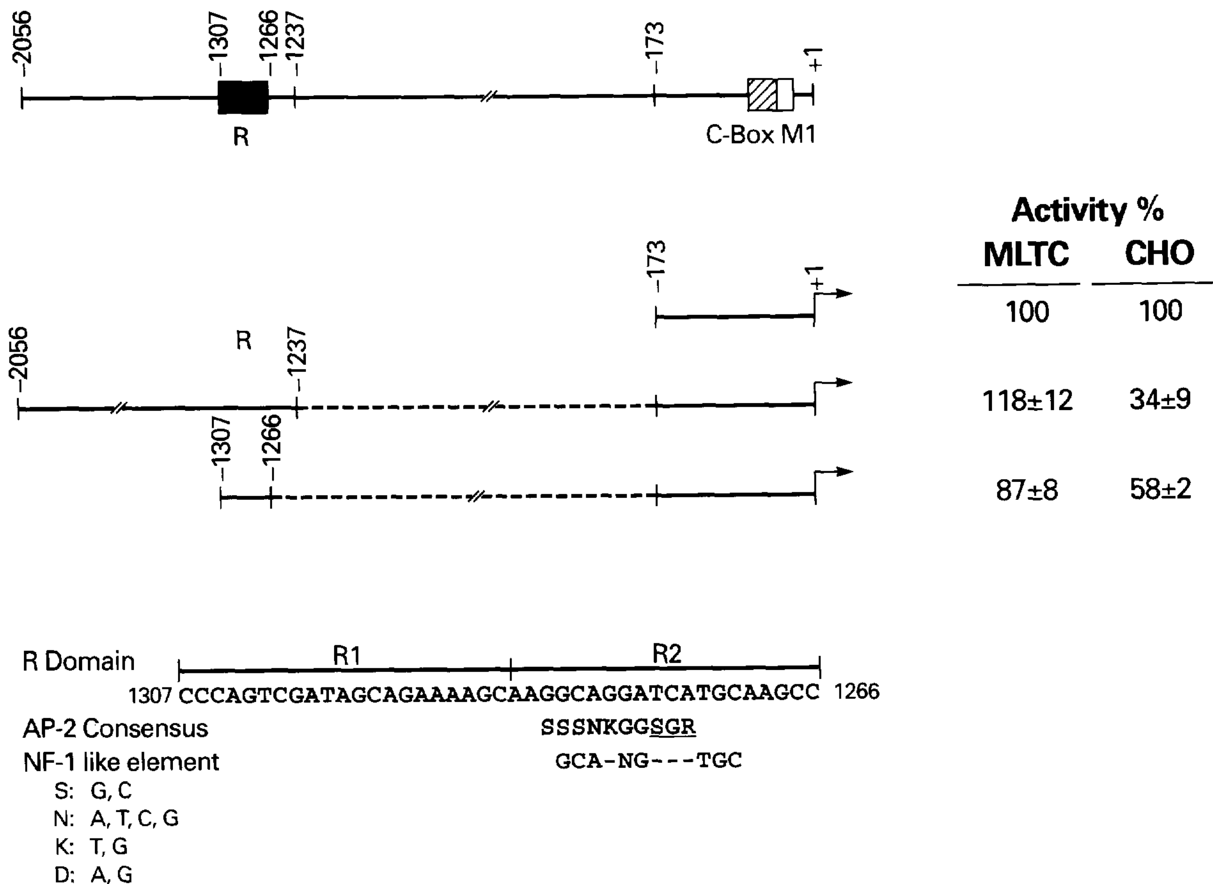

Fig. 6. Upstream inhibitory domain. Promoter activity in MLTC or CHO cells transfected with defined constructs of the upstream domain (R) adjacent to the 173 bp promoter. The nucleotide region of the R domain that binds *trans*-factor was localized to the R2 domain (below). This domain contains an AP-2 like element with three mismatches in addition to a palindrome with one mismatch that includes part of the NF-1 half element (GCA-N_6-TGC). AP-2 binds within the R2 subdomain [15, 18].

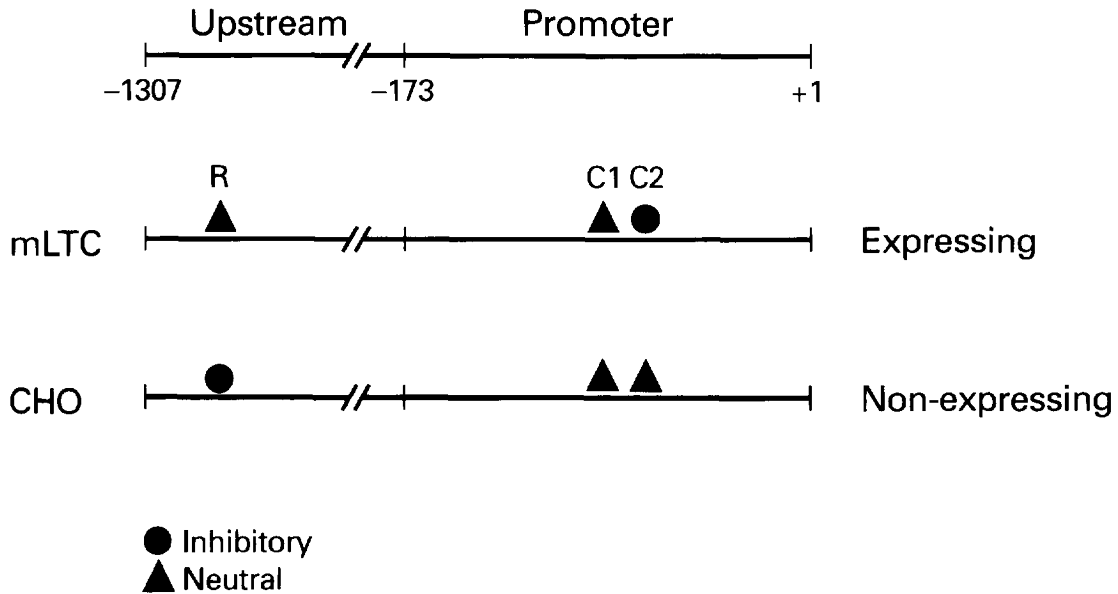

Fig. 7. Functional activity of protein binding domains of the LHR gene as deduced from mutation and deletion analysis in the expressing MLTC and non-expressing CHO cell lines [18].

resulting differences in DNA bending and transcriptional activity.

The LHR gene in expressing MLTC cells appears to be a constitutively repressed gene and requires specific activation. Two mechanisms for the reversal of inhibition in MLTC within the promoter domain that involve *trans*-factors on adjacent domains (M1/C-box) and functionally different *trans*-factors competing for the same domain (LHR AP-2 vs NF-1). Functional studies show that an AP-2-like *trans*-factor repressor preferentially binds to the upstream regulatory R domain in CHO cells, but binds to the promoter-based C-box in MLTC. This preference, most likely involving tissue-specific interference factors, may be of importance to potential mechanisms for gene activation that are only present in the expressing MLTC cell line. Binding of the repressor protein to C2 was not evident in CHO cells in functional studies, even in constructs that did not carry the upstream R domain. Repression of basal activity specifically in the nonexpressing cells was apparent only in constructs that carried the R domain upstream of the 173 bp promoter. Inhibition from this upstream domain in conjunction with removal of transfactor activation from the non-Sp1 GC box (**a**) of $Sp1_4$, and the M1 domain in non-expressing cells may play a role in silencing gene expression [18, 19]. In the expressing cell, MLTC cell, a promoter-based AP-2 repressor was evident in functional studies. This factor may compete with an NF-1-like protein for overlapping elements, to directly turn transcript initiation off or on from the promoter domain. The expressing cell, therefore, has a functional target for hormonal induction within the C-box that the non-expressing cell line does not appear to have, and this may be of importance to the cycle of LHR activation and inhibition in expressing cell types.

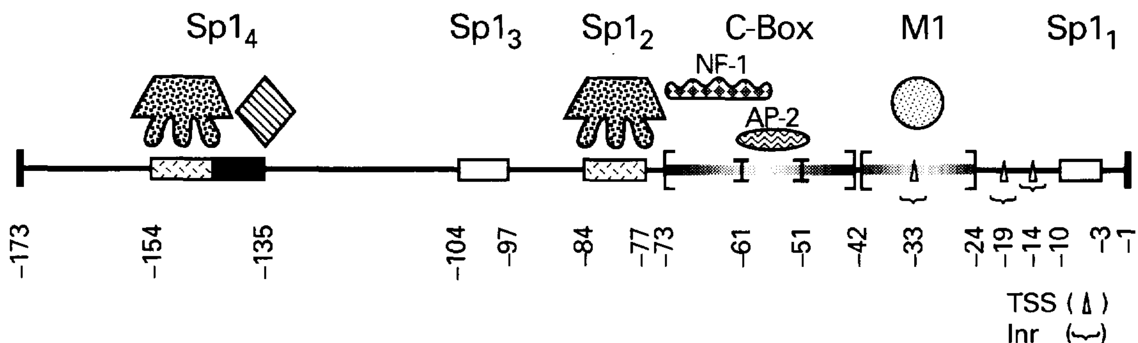

Fig. 8. Transfactors binding to promoter and upstream elements with overlapping protein binding domains on the C-Box and upstream R (R_2 subdomain) [15, 18].

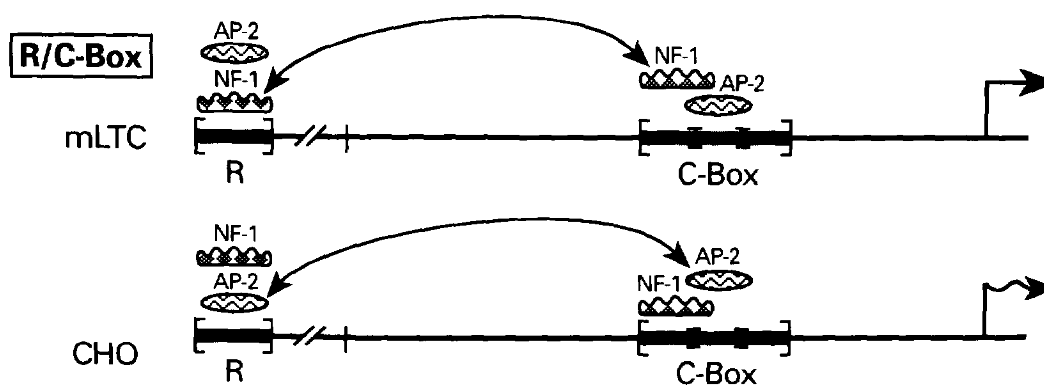

Fig. 9. Regulatory control mechanism in expressing versus non-expressing cells. Competition of AP-2 factor with an NF-1 like protein for overlapping elements at the promoter region and upstream domain contributes to regulation of promoter activity in expressing MLTC cells, and for inhibition of expression (upstream R-domain) in non-expressing cells [15, 18].

REFERENCES

1. Catt K. J. and Dufau M. L.: *Gonadotropic Hormones. Biosynthesis, Secretion Receptors and Actions in Reproductive Endocrinology* (Edited by S. S. C. Yen and R. B. Jaffe). W. B. Saunders Co., PA (1991) pp. 105–155.
2. Dufau M. L.: Endocrine regulation and communicating function of the Leydig cell. *A. Rev. Physiol.* 50 (1988) 483–508.
3. Dufau M. L.: The luteinizing hormone/human chorionic gonadotropin receptor of testis and ovary. In *Receptor Purification* (Edited by G. Litwack). Humana Press, Clifton, NJ (1990) p. 147–171.
4. Loosfelt H., Misrahi M., Atger M., Salesse R., Vu Hai-Luu Thi M. T., Jolivet A., Guiochon-Mantel A., Sar S., Jallal B., Garnier J. and Milgrom E.: Cloning and sequencing of porcine LH-hCG receptor cDNA: variants lacking transmembrane domain. *Science* 245 (1989) 525–528.
5. McFarland K. C., Sprengel R., Phillips H. S., Kohler M., Rosemblit N., Nikolics K., Segaloff D. L. and Seeburg P. H.: Lutropin-choriogonadotropin receptors: an unusual member of the G protein-coupled receptor family. *Science* 245 (1989) 494–499.
6. Tsai-Morris C-H., Buczko E., Wang W., Xie X-Z. and Dufau M. L.: Structural organization of the rat luteinizing hormone receptor gene. *J. Biol. Chem.* **266** (1991) 11,355–11,359.
7. Ji I. and Ji T. H.: Structure of the luteinizing hormone receptor gene and multiple exons of the coding sequence. *Endocrinology* 128 (1991) 2297–2308.
8. Tsai-Morris C-H., Buczko E., Wang W. and Dufau M. L.: Intronic nature of the rat luteinizing hormone receptor gene defines a soluble receptor subspecies with hormone binding activity. *J. Biol. Chem.* 265 (1990) 19,385–19,388.
9. Xie. Y. B., Wang H. and Segaloff D. L.: Extracellular domain of lutropin/choriogonadotropin receptor expressed in transfected cells binds choriogonadotropin with high affinity. *J. Biol. Chem.* 265 (1990) 21,411–21,414.
10. Minegishi T., Nakamura K., Takakura Y., Miyamoto K., Ibuki Y. and Igarashi M.: Cloning and sequencing of human LH/hCG receptor cDNA. *Biochem. Biophys. Res. Commun.* 172 (1990) 1049–1054.
11. Bernard M. P., Myers R. V. and Moyle W. R.: Cloning of rat lutropin (LH) receptor analogs lacking the soybean lectin domain. *Molec. Cell. Endocr.* **11** (1990) R19–23.
12. Hu Z. Z., Tsai-Morris C-H., Buczko E. and Dufau M. L.: Multiple and differential regulation of ovarian prolactin receptor mRNAs and their expression. *FEBS Lett.* **274** (1991) 181–184.
13. Dufau M. L., Tsai Morris C-H., Hu Z. Z., Geng Y., Xie X., Zhuang L., Zhang R. and Buczko E.: The luteinizing hormone receptor. In *Frontiers of Endocrinology. Vol. 5.: Cell and Molecular Biology of the Testis* (Edited by M. L. Dufau, A. Fabbri and A. Isidori). Ares Serono Symposia Publications, Rome, Italy (1994) pp. 51–60.
14. Hu Z-Z., Buczko E., Zhuang L. and Dufau M. L.: Sequence of the 3′-noncoding region of the luteinizing hormone receptor gene and identification of two polyadenylation domains that generate the major mRNA forms. *Biochim. Biophys. Acta* **1220** (1994) 333–337.
15. Tsai-Morris C-H., Wei W., Buczko E. and Dufau M. L.: Promoter and regulatory regions of the rat luteinizing hormone receptor gene. *J. Biol. Chem.* **268** (1993) 4447–4452.
16. Smale S. T. and Baltimore D.: The "Initiator" as transcription control element. *Cell* 57 (1989) 103–113.
17. Weiss L. and Reinberg D.: Transcription by RNA polymerase II: initiator-directed formation of transcription-competent complexes. *FASEB J.* **6** (1992) 3300–3309.
18. Tsai-Morris C-H., Gen Y., Xie X-Z., Buczko E. and Dufau M. L.: Transcriptional protein binding domains governing basal expression of the rat luteinizing hormone receptor gene. *J. Biol. Chem.* **269** (1994) 15,868–15,875.
19. Tsai-Morris C-H., Geng Y., Buczko E. and Dufau M. L.: Characterization of diverse functional elements in the upstream Sp1 domain of the rat luteinizing hormone receptor gene promoter. *J. Biol. Chem.* **270** (1995) 7487–7494.