

Cloning and Sequencing of Porcine LH-hCG Receptor cDNA: Variants Lacking Transmembrane Domain

HUGUES LOOSFELT, MICHELINE MISRAHI, MICHEL ATGER, ROLAND SALESSE, MAI TU VU HAI-LUU THI, ANDRÉ JOLIVET, ANNE GUIOCHON-MANTEL, SOKHAVUTH SAR, BAHIIJA JALLAL, JEAN GARNIER, EDWIN MILGROM*

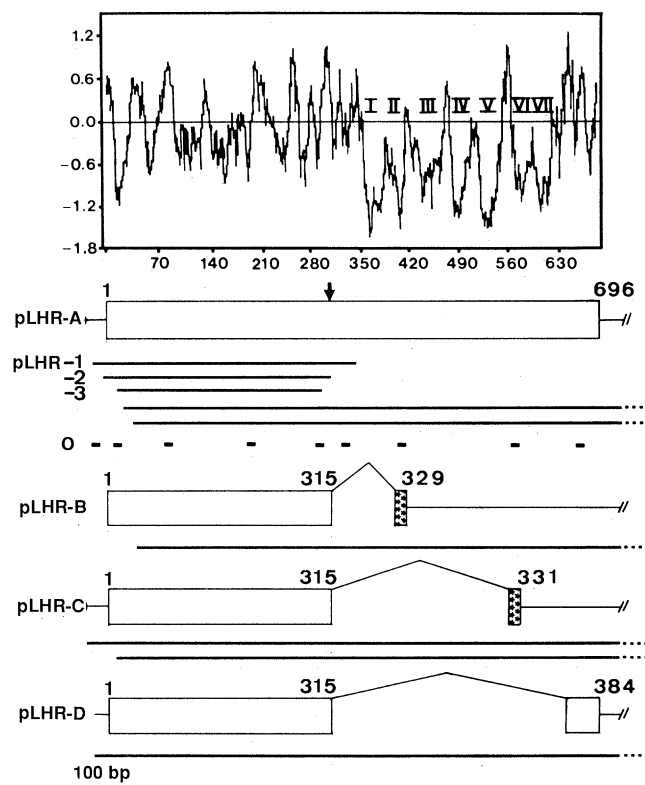
Complementary DNA clones, encoding the LH-hCG (luteinizing hormone–human choriongonadotropic hormone) receptor were isolated by screening a λ gt11 library with monoclonal antibodies. The primary structure of the protein was deduced from the DNA sequence analysis; the protein contains 696 amino acids with a putative signal peptide of 27 amino acids. Hydropathy analysis suggests the existence of seven transmembrane domains that show homology with the corresponding regions of other G protein–coupled receptors. Three other types of clones corresponding to shorter proteins were observed, in which the putative transmembrane domain was absent. These probably arose through alternative splicing. RNA blot analysis showed similar patterns in testis and ovary with a major RNA of 4700 nucleotides and several minor species. The messenger RNA was expressed in COS-7 cells, yielding a protein that bound hCG with the same affinity as the testicular receptor.

PITUITARY GLYCOPROTEINS (LUTEINIZING hormone, LH; follicle-stimulating hormone, FSH; and thyroid-stimulating hormone, TSH) form a family of closely related hormones. The structure and function of their receptors remain poorly understood (1). We have purified human choriongonadotropic (hCG)–receptor complexes from porcine testis on an Affi-Gel-10 immunomatrix to which a monoclonal antibody to β hCG had been coupled (2). Immunization of a mouse led to the preparation, and characterization of monoclonal antibodies to the receptor whose specificity was established by immunoprecipitation of hCG–receptor complexes, immunoblot, and purification of receptor by immunoaffinity chromatography (3).

Polyadenylated RNA's were isolated from porcine testes and used to prepare complementary DNA (cDNA) libraries (Fig. 1) in λ gt11 and λ gt10 vectors. Screening of the former with monoclonal antibodies to the LH-hCG receptor led to the isolation of three immunoreactive clones whose inserts were thereafter used to isolate 80 cross-hybridizing clones from the λ gt10 library. We sequenced six clones and deduced the primary structure of the complete protein (Figs. 1 and 2). The first methionine in the sequence matches the Kozak consensus (4). There is no inframe stop codon upstream, but the longest of 50 separate cDNA clones

had their 5' ends in the same region; primer extension experiments showed that we had cloned the complete messenger RNA (mRNA). The first methionine in the sequence was thus considered the initiator codon (Fig. 2). This was followed by a 27–amino acid sequence having the characteristics of a signal peptide with a cleavage site as

Fig. 1. Primary structure of the porcine LH-hCG receptor as deduced from cloned cDNA's. The full-length receptor (pLHRA) and shorter variants (pLHRB, pLHRC, and pLHRD) are indicated. The open reading frames are marked with squares (the short peptides observed after frameshift in two variants are shown as starred boxes). The divergence at amino acid 316 is shown by an arrow. The hydropathic profile (19) of the full-length protein is shown on top (the hydrophobic regions are represented by negative values). The cDNA's (from a λ gt10 library) used in the sequencing experiments are shown under each protein form and the cDNA's initially isolated from the λ gt11 library are also indicated (pLHR1-3) (20). In each case an identical structure was observed in other clones by DNA blot analysis. As four different patterns corresponding probably to different splice junctions were detected in the coding region, synthetic oligonucleotide (20 nucleotides) probes (O) complementary to each side of the junctions and to different parts of the cDNA were used to analyze 80 positive clones from the nonamplified λ gt10 library by dot blot hybridization to measure the approximate frequency of each form.



defined by Von Heijne (5). Hydropathy analysis (Fig. 1) and alignment with rhodopsin (6) (Fig. 3) suggested a possible model for the organization of the protein. A putative extracellular domain of 333 amino acids precedes a region of 266 amino acids that displays seven possible transmembrane segments. There is a 70–amino acid COOH-terminal intracellular domain. The mature protein should consist of 669 amino acids (75,025 daltons); it contains a high proportion of basic amino acids (isoelectric point, 8.5), which are localized predominantly in the COOH-terminal region. Six potential glycosylation sites (7) are found in the putative extracellular domain. Clusters of cysteines are present in the NH₂-terminal part of the protein at the junction between the putative extracellular and transmembrane domains and in the COOH-terminal region. In the COOH-terminal region there are three consensus sites for protein kinase C phosphorylation (Fig. 2) (8).

Homology searches in sequence banks (NBRF and Genpro) revealed similarities with other G protein–coupled receptors within the putative transmembrane regions (Fig. 3). The overall homology in this region is 23 percent with bovine rhodopsin (6) and 21 percent with human β 1 and β 2 adrenergic receptors (9). However, the ho-

H. Loosfelt, M. Misrahi, M. Atger, M. T. Vu Hai-Luu Thi, A. Jolivet, A. Guiochon-Mantel, S. Sar, E. Milgrom, Institut National de la Santé et de la Recherche Médicale Unité 135, Hormones et Reproduction, Hôpital de Bicêtre, 94270 Kremlin Bicêtre, France.
R. Salesse, B. Jallal, J. Garnier, Institut National de la Recherche Agronomique, Unité d'Ingénierie des Protéines, Biotechnologies, 78350 Jouy-en-Josas, France.

*To whom correspondence should be addressed.

mologies are especially focused in the putative transmembrane domains I to IV, VI, and VII (Fig. 3). Similarities are also found with the human $\alpha 1$ and $\alpha 2$ adrenergic (10), the porcine cholinergic muscarinic (11) and the bovine substance K (12) receptors, and to a lesser extent with the rat dopaminergic (13) and the human 5-hydroxytryptamine (14) receptors. Some limited regions of homology are also found with the *mas* oncogene (angiotensin receptor) (15). The long extracellular domain of LH-hCG receptor contrasts with the brevity of the corresponding region in the other receptors. It may play a role in the binding of the large glycoprotein hormone.

In addition to clones containing this large open reading frame (pLHRA), clones encoding shorter proteins were also observed.

In all cases, the NH₂-terminal part of the proteins was identical, divergence being observed at the same point (corresponding to amino acid 316 of the precursor protein). Two variants display a frameshift after the point of divergence and are thus truncated for the putative transmembrane and intracellular domains (pLHRB and pLHRC) (Fig. 1). The third variant conserves the same open reading frame after the divergence point and thus lacks the transmembrane domain but contains the putative COOH-terminal intracellular domain (pLHRD). This pattern suggests a mechanism of alternative splicing. We estimated the relative frequencies of the different variants by hybridizing 80 receptor cDNA clones from the nonamplified cDNA library with probes specific for each variant. The full-length

form (pLHRA) was the most frequent (60 percent), followed by variants D (20 percent), C (15 percent), and B (5 percent).

The shorter variants of receptor, which do not contain the putative transmembrane region, may thus be entirely extracellular, either forming quaternary receptor structures with the full-length monomer or even being secreted from the cells (as has been described for the growth hormone receptor) (16). Immunoblotting with monoclonal antibodies and protein staining after immunoaffinity purification also showed a complex receptor pattern with a major species at 85 kD and less abundant smaller forms. These results are compatible with the mRNA heterogeneity described, but molecular mass comparisons should include the glycosylation of the receptor (17).

pLHRA

ggcagcaactccgggtgcgcgcgcacctaaccagcggcgaggagcagcgtcaggctcccgccg -1	
ATG AGA CGG CGG TCC CTG GCG CTG CGG CTG CTG GCG CTG CTG CTA TTG CCA CGC CCA CTG CCG CAG ACG CTG CTC GGG GCG CCC TGC CCG GAG	96
Met Arg Arg Arg Ser Leu Ala Leu Arg Leu Leu Leu Ala Leu Leu Leu Leu Pro Pro Pro Leu Pro Gln Thr Leu Leu Gly Ala Pro Cys Pro Glu	32
CCC TGC AGC TGC CGG CCC GAC GGC GGC CTG GCG TGC CCC GGC CCG GGC GGC CTC AGC CGA CTA TCA CTC ACC TAT CTC CCT ATC AAA GTA ATC	192
Pro Cys Ser Cys Arg Pro Asp Gly Ala Leu Arg Cys Pro Gly Pro Arg Ala Gly Leu Ser Arg Leu Thr Tyr Leu Pro Ile Lys Val Ile	64
CCA TCT CAA GCT TTC AGA GGA CTT AAT GAG GTC GTA AAA ATT GAA ATC TCT CAG AGT GAT TCC CTG GAA AAG ATA GAA GCT AAT GCC TTT GAC AAC	288
Pro Ser Gln Ala Phe Arg Gly Leu Asn Glu Val Val Lys Ile Glu Ile Ser Gln Ser Asp Ser Leu Glu Lys Ile Glu Ala Asn Ala Phe Asp Asn	96
CTC CTC AAT TTG TCT GAA ATA CTG ATC CAG AAC ACC AAA AAC CTG GTG TAT ATT GAG CCT GGA GCA TTT ACA AAT CTA CCT CGG TTA AAA TAC CTG	384
Leu Leu Asn Leu Ser Leu Glu Ile Leu Ile Ser Leu Asn Thr Lys Asn Leu Val Tyr Ile Glu Pro Gly Ala Phe Thr Asn Leu Pro Arg Leu Lys Tyr Leu	128
AGC ATC TGT AAT ACA GGC ATC CGA AAG CTT CCA GAT GTT ACG AAG ATC TTC TCC TCT GAA TTT AAT TTC ATT CTG GAA ATT TGT GAT AAC TTA CAC	480
Ser Ile Cys Asn Thr Gly Ile Arg Lys Leu Pro Asp Val Thr Lys Ile Phe Ser Ser Glu Phe Asn Phe Ile Leu Glu Ile Cys Asp Asn Leu His	160
ATA ACC ACC GTA CCA GCA AAT GCT TTC CAA GGG ATG AAT AAC GAA TCC ATA ACA CTC AAA CTA TAT GGA AAT GGA TTT GAA GAA ATA CAA AGT CAT	576
Ile Thr Thr Val Pro Ala Asn Ala Phe Gln Gly Met Asn Asn Gln Ser Ile Thr Leu Lys Leu Tyr Gly Asn Gly Phe Glu Glu Ile Gln Ser Ile	192
GCC TTC AAT GGG ACA ACG CTG ATT TCC CTG GAG CTG AAG GAG AAT GCA CAC CTG AAG AAG ATG CAC AAT GAC GCC TTC CGA GGG GCC AGA GGG CCC	672
Ala Phe Asn Gly Thr Thr Leu Ile Ser Leu Glu Leu Lys Glu Asn Ala His Leu Lys Lys Met His Asn Asp Ala Phe Arg Gly Ala Arg Gly Pro	224
AGC ATC TTG GAT ATT TCT TCC ACT AAA CTA CAG GCC CTG CCT AGT TAT GGG CTG GAG TCC ATT CAG ACG CTA ATT GCC ACA TCA TCC TAT TCT CTG	768
Ser Ile Leu Asn Ser Ile Ser Ile Phe Lys Asn Phe Ser Lys Gln Cys Glu Ser Thr Ala Arg Arg Pro Asn Asn Glu Thr Leu Tyr Ser Ser Tyr Ser Leu	256
AAA AAA CTG CCA TCA AGA GAA AAA TTT ACC AAT CTC CTA GAT GCC ACA TTG ACT TAC CCC AGC CAC TGC TGT GCT TTT AGA AAC CTG CCA ACA AAA	864
Lys Lys Leu Pro Ser Arg Glu Lys Phe Thr Asn Leu Leu Asp Ala Thr Leu Thr Tyr Pro Ser His Cys Cys Ala Phe Arg Asn Leu Pro Thr Lys	288
GAG CAG AAT TTT TCA TTT TCC ATT TTT AAA AAC TTT TCT AAA CAA TGT GAA AGC ACA GCA AGG AGA CCA AAT AAT GAA ACA CTT TAT TCT GCC ATC	960
Glu Gln Asn Phe Ser Ile Phe Lys Asn Phe Ser Lys Gln Cys Glu Ser Thr Ala Arg Arg Pro Asn Asn Glu Thr Leu Tyr Ser Ala Ile	320
TTT GCT GAG AGT GAA CTG AGT GAC TGG GAT TAT GAC TAT GGT TTC TGC TCA CCC AAG ACA CTC CAA TGT GCT CCC GAA CCA GAT GCT TTT AAC CCC	1056
Phe Ala Glu Ser Glu Leu Ser Asp Trp Asp Tyr Asp Tyr Gly Phe Cys Ser Pro Lys Thr Leu Gln Cys Ala Pro Glu Pro Asp Ala Phe Asn Pro	352
TGT GAA GAT ATT ATG GGC TAT GAC TTC CTT AGA GTT CTG ATT TGG CTG ATT AAT ATT CTA GCC ATT ATG GGA AAC GTG ACT GTC CTC TTT GTT CTC	1152
Cys Glu Asp Ile Met Gly Tyr Asp Phe Leu Arg Val Leu Ile Trp Leu Ile Asn Ile Leu Ala Ile Met Gly Asn Val Thr Val Leu Phe Val Leu	384
CTG ACC AGT CAT TAT AAA CTG ACA GTG CCT CGT TTC CTC ATG TGC AAT CTC TCC TTT GCA GAC TTC TGC ATG GGG CTC TAC CTG CTA CTC ATT GCC	1248
Leu Thr Ser His Tyr Lys Leu Thr Val Pro Arg Phe Leu Met Cys Asn Leu Ser Phe Ala Asp Phe Cys Met Gly Leu Tyr Leu Leu Ile Ala	416
TCA GTT GAT GCC CAA ACC AAA GGC CAG TAT TAT AAC CAC GCC ATG GAG TGG CAG ACA GGG AAT GGG TGT AGT GTT GGT GGC TTT TTC ACT GTA TTT	1344
Ser Val Asp Ala Gln Thr Lys Gly Gln Tyr Tyr Asn His Ala Ile Asp Trp Gln Thr Gly Asn Gly Cys Ser Val Ala Gly Phe Phe Thr Val Phe	448
GCA AGT GAA CTT TCT GTC TAC ACC CTC ACA GTC ATC ACA CTA GAA AGA TGG CAT ACC ATC ACC TAT GCT ATT CAG CTG GAC CAA AAG CTA CGC TTA	1440
Ala Ser Glu Leu Ser Val Tyr Thr Leu Thr Val Ile Thr Leu Glu Arg Trp His Thr Ile Thr Ala Ile Gln Leu Asp Gln Lys Leu Arg Leu	480
AGA CAT GCC ATT CCA ATT ATG CTT GGA GGG TGG CTC TTT TCT ACA CTA ATT GCC ATG TTG COT CTT GTG GGT GTC AGC AGT TAC ATG AAG GTC AGC	1536
Arg His Ala Ile Pro Ile Met Leu Gly Gly Trp Leu Phe Ser Thr Leu Ile Ala Met Leu Pro Leu Val Gly Val Ser Ser Tyr Met Lys Val Ser	512
ATT TGC CTC CCC ATG GAT GTG GAA ACC ACT CTC TCA CAG GTC TAC ATA TTA ACC ATC CTG ATC CTC AAT GTG GTG GCC TTC ATC ATC ATT TGT GCT	1632
Ile Cys Leu Pro Met Asp Val Glu Thr Thr Leu Ser Gln Val Tyr Ile Leu Thr Ile Leu Lys Leu Asn Val Val Ala Phe Ile Ile Ile Cys Ala	544
TGC TAC ATT AAA ATT TAT TTT GCA GTT CAA AAT CCA GAG CTG ATG GCT ACC AAC AAA GAC ACA AAG ATT GCT AAG AAA ATG GCA GTC CTC ATC TTC	1728
Cys Tyr Ile Lys Ile Tyr Phe Ala Val Gln Asn Pro Glu Leu Met Ala Thr Asn Lys Asp Thr Lys Ile Ala Lys Lys Met Ala Val Leu Ile Phe	576
ACT GAT TTC ACC TGC ATG GCA CCG ATC TCT TTC TTT GCC ATC TCA GCT GCC TTA AAA GTG CCC CTT ATT ACA GTA ACG AAC TCT AAG GTA TTA CTG	1824
Thr Asp Phe Thr Cys Met Ala Pro Ile Ser Phe Phe Ala Ile Ser Ala Ala Leu Lys Val Pro Leu Ile Thr Val Thr Asn Ser Lys Val Leu Leu	608
GTT CTT TTT TAT CCT GTC AAT TCT TGT GCC AAT CCG TTT CTA TAC GCA ATT TTC ACA AAG GCA TTC CGA AGG GAT TTC TTT CTG TTG CTG AGC AAA	1920
Val Leu Phe Tyr Pro Val Asn Ser Cys Ala Asn Pro Phe Leu Tyr Ala Ile Phe Thr Lys Ala Phe Arg Arg Asp Phe Phe Leu Leu Leu Ser Lys	640
TCT GGC TGC TGT AAA CAT CAA GCC GAA CTT TAT AGA CCG AAG GAC TTT TCA GCT TAC TGC AAA AAT GGC TTC ACT GGA TCA AAT AAG CCT TCC CGG	2016
Ser Gly Cys Cys Lys His Gln Ala Glu Leu Tyr Arg Arg Lys Thr Ser Thr Lys Lys Asn Gly Phe Thr Gly Ser Asn Lys Pro Ser Arg	672
TCC ACA TTG AAG TTG ACC ACA TTA CAA TGT CAG TAT TCC ACT GTC ATG GAC AAG ACT TGC TAT AAA GAC TGT taactgtgtatgagtaaccacataatttaa	2119
Ser Thr Leu Lys Leu Thr Thr Leu Gln Cys Gln Tyr Ser Thr Val Met Asp Lys Thr Cys Tyr Lys Asp Cys	696

pLHRB

ACA CTA CTT CTG CAT GGG GCT CTA CCT GCT ACT CAT TGC CTC AGT TGA	987
Thr Leu Leu Leu His Gly Ala Leu Pro Ala Thr His Cys Leu Ser	329
ACA CTT TCA AAA TCC AGA GCT GAT GGC TAC CAA CAA AGA CAC AAA GAT TGC TAA	993
Thr Leu Ser Lys Ser Arg Ala Asp Gly Tyr Gln Gln Arg His Lys Asp Cys	331
ACA CTG GCA TTC CGA AGG GAT TTC TTT CTG TTG CTG ... AAA GAC TGT TAA	1152
Thr Leu Ala Phe Arg Arg Asp Phe Phe Leu Leu ... Lys Asp Cys	384

Fig. 2. Sequence of the porcine LH-hCG receptor cDNA and of the deduced protein. (▽) Point of divergence between full-length and shorter forms of receptor. (▼) Points of junction with the main sequence in the various shorter forms. The sequence of the shorter variants is shown from amino acid 315 (preceding the divergence). For pLHRD, apart from a single amino acid change at the junction, the same amino acid sequence is observed from residue 317 to the end (matching exactly residues 629 to 696 of LHRA). N-linked glycosylation sites (solid lines) and protein kinase C consensus phosphorylation sites (dotted lines) are underlined.

Fig. 3. Homologies between LH-hCG receptor and other G protein-coupled receptors. The bovine rhodopsin (bRHOD), and human β_1 (hBETA1), β_2 (hBETA2), α_1 (hALPHA1), and α_2 (hALPHA2) adrenergic, porcine muscarinic cholinergic (pm3ACH) and bovine substance K (bSK) receptors were compared to full-length LH receptor. Amino acids identical in the latter and in at least one of the other receptors are indicated by brackets.

	I	II	III	IV	V	VI	VII
pLH/hCG	356						
bRHOD	IMGYDEIRVIMINILAIMGNVIVLVLTSMKI--TPR--FLMONTSFDFCMGIYLLIASDDQTKGMYNHAID--G--G--CSVAGFFVFASELE						
hBETA1	PWQFSMLAAYMFLIMLGFPINFLTMVTVO--HKKLRIPLNLY--LLNLAVADLFMVEGGFTTLYTSLH--G--G--EVEFGTCNLEGGFATLGGC						
hBETA2	QWTAG--MGLIMAILVILVAGNVIMVAIAK--TPRLQTLNIF--IMSLASADLVMGILLVVPFGATIVVW--G--G--WEYGSFCELTWSDVLCVTA						
hALPHA1	VWVVG--MGIVMSLIVLAVIFGNVIMVAIAK--FERLOTNTNIF--ITSLACADLVMGILLVVPFGAAILH--G--G--WTEGFWCEFTWSDVLCVTA						
hALPHA2	TRAI--VGLVLGAFILFAIVGNVIMVLSVAC--NRHLRPTNIF--IVNLADLLLSFTVLPFSATLEVL--G--G--WVIGRIFCDIWAARDVLCCTAS						
pm3ACH	QVTLT--IVCLAGLMLITVFGNVIMVAIAK--SRALKAPONIF--LVSLASADLVATLVIPESLANEVM--G--G--WYFGKTCWEIYALDLCTCS						
bSK	PWQVAFIGITTGLLSLATVTGNLLVLSFKV--NTELTNNYF--LLSLACADLVITFSMNLYTLYLLM--G--G--WALGTACDLWLALDYVASNAS						
	LWTAAYLA--LVLVAVMGNATVITWILDA--BORMRTNIN--IVNLADLVMAAFNAAFNEV--G--G--WYFGRAHCYFQNLPIITAMFVS						
pLH/hCG							
bRHOD	VYTLTVILERWHITITATOLDKIRLEHATPMLGNDSTLIAMPLVGVSSYMKVSI--CL--BMDVPHLSQVHITITLUN--VMAFITICA						
hBETA1	LWSILMLATERVYVVCCKPMS--NREGENHAIMGVAFITVMA--LACAPFLVGSRIPEG--MQCSCGIDYTPH--NNESFVIMFVVEHIIPLIMIFF						
hBETA2	IEITLVINLPRYLATISPFYQSLITRARGVLCVMAISALVSFLPILMHWRRAESDEAR--RCYND--KCCDFVTNRATIASSVVSFYPLQIMAF						
hALPHA1	IEITLVINLPRYLATISPFYQSLITRARGVLCVMAISALVSFLPILMHWRRAESDEAR--RCYND--KCCDFVTNRATIASSVVSFYPLQIMAF						
hALPHA2	IEITLVINLPRYLATISPFYQSLITRARGVLCVMAISALVSFLPILMHWRRAESDEAR--RCYND--KCCDFVTNRATIASSVVSFYPLQIMAF						
pm3ACH	IEITLVINLPRYLATISPFYQSLITRARGVLCVMAISALVSFLPILMHWRRAESDEAR--RCYND--KCCDFVTNRATIASSVVSFYPLQIMAF						
bSK	IEITLVINLPRYLATISPFYQSLITRARGVLCVMAISALVSFLPILMHWRRAESDEAR--RCYND--KCCDFVTNRATIASSVVSFYPLQIMAF						
pLH/hCG							
bRHOD	CYINITYFAYONPELMATNITKAKKMLITDFT--CHAFIS--FAISALPLITVTSNKV--LIVLEY--FVNSCANPFLYAFITKAFRR--LILLKS						
hBETA1	CYQQLMTEVKE--(12)--HAEKEVIRMTIMVIAFLICWLPY--LAGVAFYIFTHQGSDFGIFMTIPAFFAKTSAYVNPVIMMNEF--NCMVTILCCG						
hBETA2	VYLRVFEAKR--(64)--REKKAUK--TLGITMGVFTLCWLP--FFLANVVKAFHREL--VPDR--LVEENWLGYNANSAFNILIN--CRSPDFRKAFOGLLCCA						
hALPHA1	VYLRVFEAKR--(38)--REKKAUK--TLGITMGVFTLCWLP--FFIVNIVHVIQDNIL--IRK--EVYIILNWIWVNSGFNPLIN--CRSPDFRKAFOGLLCCA						
hALPHA2	VYLRVFEAKR--(55)--REKKAUK--TLGITMGVFTLCWLP--FFIALPLGSLSFTLKPPD--AVFKVFWLGYNSGLNIN--IMPSSHFKRAFRMILGCO						
pm3ACH	VYLRVFEAKR--(141)--NLEKRETFVLAVVIGVFWVWLP--FFFTYTLTAVSCS--VPR--TLKKEFWFGYVNSGLNIN--IMPSSHFKRAFRMILGCO						
bSK	VYLRVFEAKR--(141)--NLEKRETFVLAVVIGVFWVWLP--FFFTYTLTAVSCS--VPR--TLKKEFWFGYVNSGLNIN--IMPSSHFKRAFRMILGCO						
	LYWRIYRETN--(141)--REKKAUK--TLGITMGVFTLCWLP--FFFTYTLTAVSCS--VPR--TLKKEFWFGYVNSGLNIN--IMPSSHFKRAFRMILGCO						
	ANVIGLTLWR--(19)--QAKKFKVKTMLVWVITATC--LWLYHYLIGLTFQEDIYCHKFIQVYLALFWLAMSSTMYNI--IVCCLNHRFRSCRLAFRC						

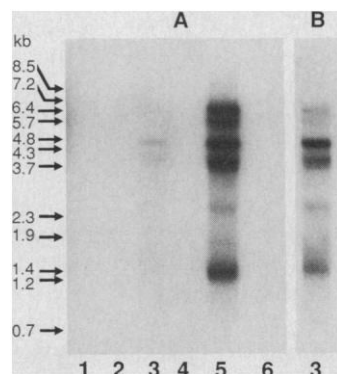


Fig. 4. RNA blot analysis of porcine LH-hCG receptor mRNA in different tissues. (Lane 1) Liver; (lane 2) intestine; (lane 3) ovary; (lane 4) lung; (lane 5) testis; and (lane 6) muscle. Blots were exposed for 20 hours (A) or 1 week (B) to "no screen" x-ray films. Sizes of DNA markers (kilobases) are indicated on the left. Polyadenylated RNA's were prepared as described (Fig. 1). Each lane contained 20 μ g of polyadenylated RNA. Hybridization was performed with the random-primed 32 P-labeled 1-kb insert of the pLHR1 clone (6×10^6 dpm/ng, 0.5×10^6 dpm/ml). Receptor concentration (as measured by 125 I-labeled hCG binding) in testes and ovaries was 200 and 40 fmol of protein per milligram, respectively. Receptor was not detected in the other organs.

Blot analyses of polyadenylated RNA's from various organs showed the presence in testes and ovaries of the same mRNA species (Fig. 4): a major mRNA of 4700 nucleotides and minor species of 6700 (abundant in testis), 5800, 4000, 2600, and 1400 nucleotides. No receptor mRNA was detected in liver, intestine, lung, and muscle. The receptor mRNA concentrations, as estimated from the RNA blot, paralleled hormone binding assays in the same tissues (Fig. 4).

The cDNA containing the longest open reading frame was inserted into the

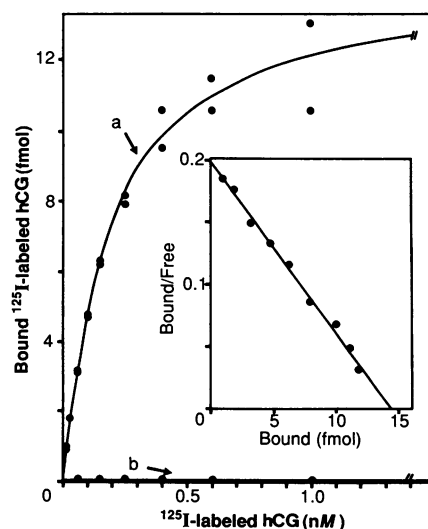


Fig. 5. Expression of the porcine LH-hCG receptor cDNA in COS-7 cells. Binding of 125 I-labeled hCG to membranes (25). (Curve a) Cells transfected with a vector encoding LH-hCG receptor. (Curve b) Cells transfected with a vector containing inverted cDNA. (Inset) Scatchard plot of the binding assay ($K_d = 1.8 \times 10^{-10}$ M).

pKSV10 expression vector, and the resulting DNA was transfected into COS-7 cells (Fig. 5). This led to the appearance of the cell membrane of a hCG-binding protein with an affinity ($K_d = 1.8 \times 10^{-10}$ M) similar to that displayed by the receptor in porcine Leydig cells (18).

Our data suggest that the protein encoded by the cloned cDNA contains all the information necessary to bind the hormone. The protein also has the characteristic transmembrane structure of adenylate cyclase-modulating receptors. However, further experiments are necessary to establish whether the protein can activate adenylate cyclase by itself and promote steroidogenesis regula-

tion. Moreover, receptor immunoaffinity purification experiments provided no evidence for the existence of any other protein component (3). The role of the relatively abundant truncated forms of the receptor is not understood at present, and it is not known whether the LH-hCG receptor is physiologically active as a monomer or as an oligomer. The cloning of the cDNA encoding the porcine LH-hCG receptor should now allow the isolation of the corresponding human receptor cDNA and possibly that of the receptors binding the other members of the pituitary glycoprotein hormone family.

REFERENCES AND NOTES

1. S. Kusuda and M. L. Dufau, *J. Biol. Chem.* **261**, 16161 (1986); K. P. Keinanen, S. Kellokumpu, M. K. Metsikko, H. J. Rajaniemi, *ibid.* **262**, 7920 (1987); H. T. Sojar and O. P. Bahl, *ibid.* **264**, 2552 (1989); M. Ascoli and D. L. Segaloff, *Endocrinol. Rev.* **10**, 27 (1989).
2. B. Jallat *et al.*, in preparation.
3. M. T. Vu Hai-Luu Thi *et al.*, in preparation.
4. M. Kozak, *Nature* **308**, 241 (1984).
5. G. Von Heijne, *Nucleic Acids Res.* **14**, 4683 (1986).
6. J. Nathans and D. S. Hogness, *Cell* **34**, 807 (1983).
7. S. C. Hubbard and R. J. Ivatt, *Annu. Rev. Biochem.* **50**, 555 (1981).
8. J. R. Wodget, K. L. Gould, T. Hunter, *Eur. J. Biochem.* **161**, 177 (1986).
9. T. Frielle *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7920 (1987); B. K. Kobilka *et al.*, *ibid.*, p. 46.
10. S. Cotecchia *et al.*, *ibid.* **85**, 7159 (1988); B. K. Kobilka *et al.*, *Science* **238**, 650 (1987).
11. T. Kubo *et al.*, *Nature* **323**, 411 (1986).
12. Y. Masu *et al.*, *ibid.* **329**, 836 (1987).
13. J. R. Burnzow *et al.*, *ibid.* **336**, 783 (1988).
14. A. Fargin *et al.*, *ibid.* **335**, 358 (1988); D. Julius, A. B. MacDermott, R. Axel, T. M. Jessel, *Science* **241**, 558 (1988); D. B. Pritchett *et al.*, *EMBO J.* **7**, 4135 (1988).
15. T. R. Jackson, L. A. C. Blair, J. Marshall, M. Goedert, M. R. Hanley, *Nature* **335**, 437 (1988).
16. D. W. Leung *et al.*, *ibid.* **330**, 537 (1987).
17. K. P. Keinanen, *Biochem. J.* **256**, 719 (1988).
18. N. Genty, R. Salesse, J. Garnier, *Biol. Cell.* **59**, 129 (1987).
19. T. P. Hopp and K. R. Woods, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 3824 (1981).

20. Polyadenylated RNA's were prepared (21) from piglet testes (10 to 20 days old). Random-primed cDNA synthesis, size selection (0.5 to 2 kb), and cloning into Agt11 vector were as described (22). The expression library (3×10^6 clones) was screened (23) with a mixture of ten monoclonal antibodies to the porcine LH-hCG receptor (3). Three positive clones that cross-hybridized were selected (pLHR 1 to 3) and sequenced. The longer insert (a 1-kb Eco RI fragment from pLHR1) was ^{32}P -labeled by random priming and was used to probe 2×10^6 clones from a second size-selected (2 to 6 kb) cDNA library in Agt10 vector. Six overlapping clones from the 400 detected signals (selected for length) were used for sequencing (24).
21. H. Loosfelt, F. Logeat, M. T. Vu Hai, E. Milgrom, *J. Biol. Chem.* **259**, 14196 (1984).
22. H. Loosfelt *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 9045 (1986).
23. T. V. Huynh, R. A. Young, R. W. Davis, in *DNA Cloning: A Practical Approach*, D. M. Glover, Ed. (IRL, Oxford, 1985), vol. 1, p. 98.
24. M. Misrahi *et al.*, *Biochem. Biophys. Res. Commun.* **143**, 740 (1987).
25. The full-length open reading frame of the receptor (nucleotides -11 to 2113, see sequence in Fig. 2) was inserted into the Bgl II site of the pKSV10 vector (Pharmacia). COS-7 cells were grown and transfected (26) with 10 μg of vector DNA per 10^6 cells. After 42 hours of transient expression, cell membranes were isolated (27). Radioligand binding assays were carried out in duplicate: membranes (1.5 mg of protein per milliliter) were incubated in 0.4 ml of phosphate-buffered saline containing 0.1% sodium azide and various amounts of ^{125}I -labeled hCG (0.5 $\mu\text{Ci}/\text{pmol}$), for 12 hours at 37°C . The mixture was centrifuged at $10,000g$ and the pellets were washed once with 1 ml of buffer; bound radioactivity was then measured. The nonspecific binding was determined by parallel incubations with unlabeled 2.10^{-7}M hCG and this amount was subtracted from total binding.
26. A. Guiochon-Mantel *et al.*, *Nature* **336**, 695 (1988).
27. V. Rebois, *J. Cell. Biol.* **94**, 70 (1982).
28. We thank P. Dessen for help in computer analysis and D. Grebert for technical assistance.

23 May 1989; accepted 23 June 1989

Domain Separation in the Activation of Glycogen Phosphorylase *a*

E. J. GOLDSMITH, S. R. SPRANG, R. HAMLIN, N.-H. XUONG, R. J. FLETTERICK

The crystal structure of glycogen phosphorylase *a* complexed with its substrates, orthophosphate and maltopentaose, has been determined and refined at a resolution of 2.8 angstroms. With oligosaccharide bound at the glycogen storage site, the phosphate ion binds at the catalytic site and causes the regulatory and catalytic domains to separate with the loss of stabilizing interactions between them. Homotropic cooperativity between the active sites of the allosteric dimer results from rearrangements in isologous contacts between symmetry-related helices in the subunit interface. The conformational changes in the core of the interface are correlated with those observed on covalent activation by phosphorylation at Ser¹⁴ (phosphorylase *b* $\rightarrow$ *a*).

GLYCOGEN PHOSPHORYLASE (GP) (E.C. 4.2.1.1) is a cooperative homodimer that catalyzes the degradative phosphorolysis of glycogen. The enzyme is regulated by covalent phosphorylation (1) and by substrates and effectors (2). The substrates phosphate, glucose-1-phosphate (G1P), and glycogen bind cooperatively (3) and, in terms of the two-state model (4), drive the conformational equilibrium from the catalytically inactive T-state to the active R-state conformation. Of special interest are the conformational changes required to bind these substrates and the mechanisms by which the allosteric effectors promote these changes. Comparison of the high-resolution crystal structures of the dephosphorylated (GP_b) and phosphorylated (GP_a) enzymes (5), the latter under glucose inhibition, has shown that phosphorylation of Ser¹⁴ results in the formation of extensive subunit contacts on the regulatory face of

the molecule. Similar changes are induced by the activator adenosine monophosphate (AMP) (6). However, the active-site-bound glucose in GP_a crystals blocks structural changes at the catalytic site that might otherwise occur in the free enzyme on phosphorylation. Previous attempts to grow crystals of GP_a with substrates bound or in the absence of glucose have been unsuccessful. Madsen *et al.* (7) found that fresh crystals of GP_a tend to disintegrate when exposed to a combination of AMP, oligosaccharide, and G1P, although such crystals can reanneal to form a stable lattice with expanded unit cell dimensions. In contrast, crystals of GP_a that have become naturally cross-linked while aging maintain their integrity in the presence of substrates and activators and diffract to moderate resolution.

To probe the active conformation of phosphorylase *a*, we substituted glucose in crystals of GP_a with orthophosphate and maltopentaose by diffusion. The refined 2.8 Å resolution structure contains phosphate ion bound in the catalytic site and maltopentaose bound only in the storage-activation site. Together, the two substrates induce local conformational changes in the catalytic site and a domain separation, as well as rearrangements at the subunit interface that are not observed when either alone is pre-

sent (8, 9). However, natural cross-linking and lattice forces may restrain the enzyme from undergoing a complete transition to the R state.

After diffusion of the substrates orthophosphate and maltopentaose into crystals of GP_a, the volume of the tetragonal unit cell (Table 1) increases by 28,100 Å³ (3), primarily as a result of a 1.7 Å expansion along the crystallographic *c*-axis, and the resolution of observable diffraction decreases from 2.0 to 2.8 Å (10). Diffraction data were measured on the Mark II multiwire area detector (11). Crystals containing orthophosphate and maltopentaose show substantial nonisomorphism with respect to the parent crystals, with an average change of 43% in structure-factor amplitudes (8). The structure was determined by a sequence of rigid body and restrained group

Table 1. Crystal data and refinement parameters for substrate-inhibited GP_a (space group $P4_32_12$). For comparison, data are also shown for glucose-inhibited GP_a. The scaling *R*-factor between data sets is 0.36; $R_{\text{sym}} = \{\sum_{hkl} \sum_n (I_{hkl} - I_{hkl}) / I_{hkl}\} / N_{hkl}$, where I_{hkl} is the *n*th measurement of I_{hkl} ; $R_{\text{cryst}} = \sum_{hkl} [|F(\text{obs})_{hkl}| - |F(\text{calc})_{hkl}|] / |F(\text{obs})_{hkl}|$, where $|F(\text{obs})|$ and $|F(\text{calc})|$ are the observed and calculated structure factor amplitudes, respectively.

Parameter	GP _a	
	Activated	Inhibited
<i>Crystal data</i>		
Unit cell dimensions		
<i>a</i> (Å)	128.5	128.4
<i>c</i> (Å)	118.1	116.4
Resolution (Å)	2.8	2.1
Reflections	23,035	35,231
<i>R</i> _{sym}	0.06	0.05
<i>Refinement parameters</i>		
Protein atoms	6695	6624
Δ bond* (Å)	0.013	0.015
Δ angle* (Å)	0.050	0.050
Resolution (Å)	50 to 2.8	50 to 2.1
Reflections	21,753	34,792
<i>R</i> _{cryst}	0.19	0.16

*Applied restraints were 0.01 Å on bonds and 0.03 Å on angles.

E. J. Goldsmith, Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX 75235.

S. R. Sprang, Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, TX 75235.

R. Hamlin and N.-H. Xuong, Departments of Physics, University of California San Diego, La Jolla, CA 92093.

R. J. Fletterick, Department of Biochemistry and Biophysics, University of California San Francisco, San Francisco, CA 94143.