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, BAHIJA JALLAL, JEAN GARNIER, EDWIN MILGROM*

Complementary DNA clones, encoding the LH-hCG (luteinizing hormone–human choriogonadotropic hormone) receptor were isolated by screening a λgt11 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 choriogonadotropic (hCG)–receptor complexes from porcine testis on an Affi-Gel-10 immunomatrix to which a monoclonal antibody to βhCG 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 λgt11 and λgt10 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 λgt10 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 condon (Fig. 2). This was followed by a 27–amino acid sequence having the characteristics of a signal peptide with a cleavage site as 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 NH_2-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 β1 and β2 adrenergic receptors (*9*). However, the ho-

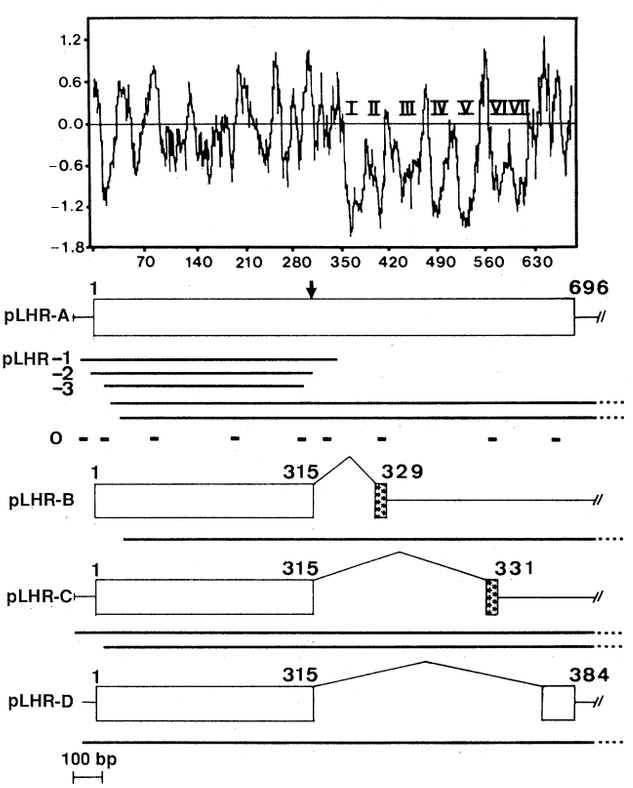

Fig. 1. Primary structure of the porcine LH-hCG receptor as deduced from cloned cDNA's. The full-length receptor (pLHRA) and shorter variants (pLHRB, pLHRC, and pLHRD) are indicated. The open reading frames are marked with squares (the short peptides observed after frameshift in two variants are shown 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 λgt10 library) used in the sequencing experiments are shown under each protein form and the cDNA's initially isolated from the λgt11 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 λgt10 library by dot blot hybridization to measure the approximate frequency of each form.

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 Bicetre, 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 α1 and α2 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 NH_2-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*).

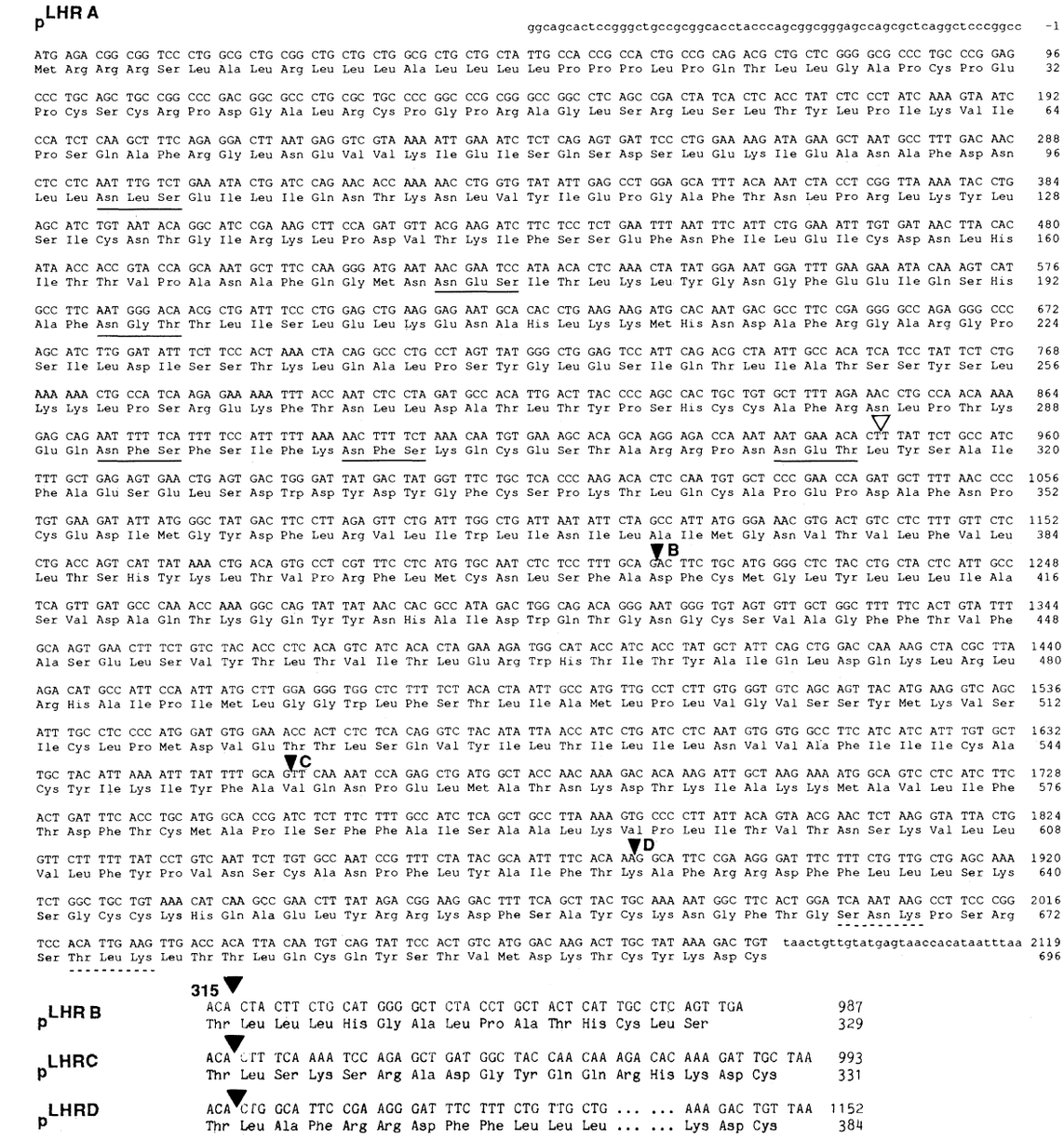

Fig. 2. Sequence of the porcine LH-hCG receptor cDNA and of the deduced protein. (▽) Point of divergence between full-length and shorter forms of receptor. (▼) Points of junction with the main sequence in the various shorter forms. The sequence of the shorter variants is shown from amino acid 315 (preceding 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.

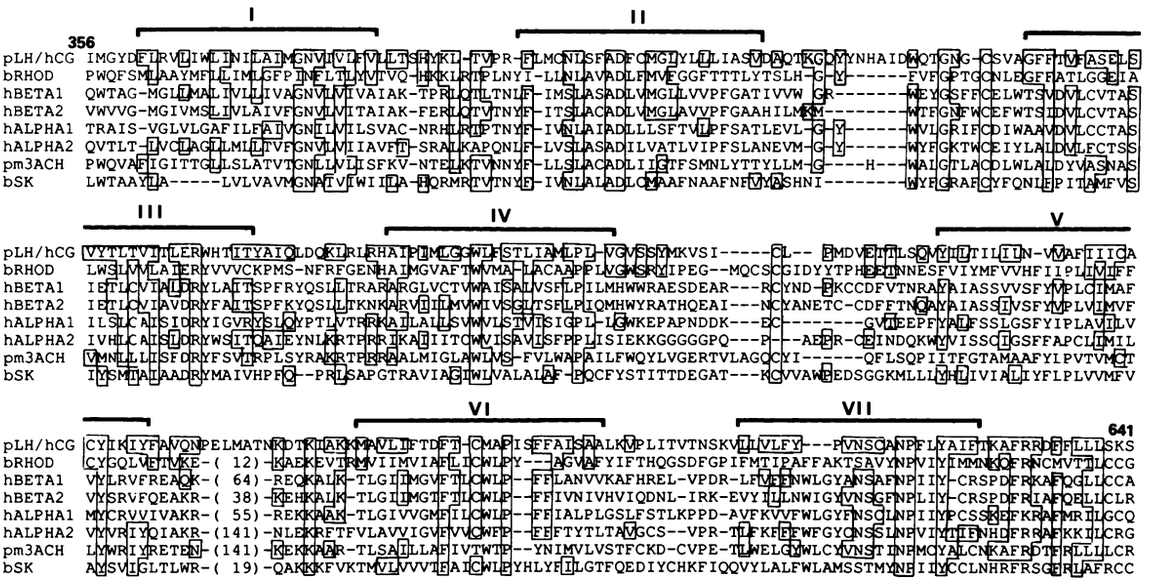

Fig. 3. Homologies between LH-hCG receptor and other G protein–coupled receptors. The bovine rhodopsin (bRHOD), and human β_1(hBETA1), β_2(hBETA2), α_1(hALPHA1), and α_2(hALPHA2) adrenergic, porcine muscarinic cholinergic (pm3-ACH) and bovine substance K (bSK) receptors were compared to full-length LH receptor. Amino acids identical in the latter and in at least one of the other receptors are squares. The putative transmembrane domains (I to VII) are indicated by brackets.

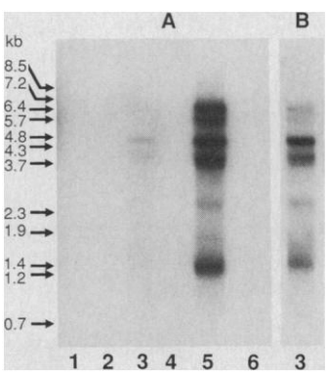

Fig. 4. RNA blot analysis of porcine LH-hCG receptor mRNA in different tissues. (Lane 1) Liver; (lane 2) intestine; (lane 3) ovary; (lane 4) lung; (lane 5) testis; and (lane 6) muscle. Blots were exposed for 20 hours (**A**) or 1 week (**B**) to "no screen" x-ray films. Sizes of DNA markers (kilobases) are indicated on the left. Polyadenylated RNA's were prepared as described (Fig. 1). Each lane contained 20 μg 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.

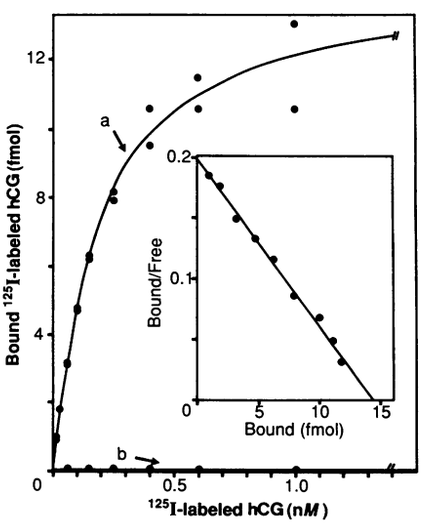

Fig. 5. Expression of the porcine LH-hCG receptor cDNA in COS-7 cells. Binding of ^{125}I-labeled hCG to membranes (*25*). (Curve a) Cells transfected with a vector encoding LH-hCG receptor. (Curve b) Cells transfected with a vector containing inverted cDNA. (Inset) Scatchard plot of the binding assay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

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 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 regulation. 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. Jallal *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. Bunzow *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 λgt11 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 λgt10 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 μCi/pmol), for 12 hours at 37°C. The mixture was centrifuged at 10,000*g* 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 oligosaccaride 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 Ser^{14} (phosphorylase *b* → *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 Ser^{14} 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 present (*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 Å (*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; $R_{sym} = \{\Sigma_{hkl}\Sigma_n[(I^n_{hkl} - I_{hkl})/I_{hkl}]\}/N_{hkl}$, where I^n_{hkl} is the *n*th measurement of I_{hkl}; $R_{cryst} = \Sigma_{hkl}\ [||F(obs)_{hkl}| - |F(calc)_{hkl}||]/|F(obs)_{hkl}|$, where $|F(obs)|$ and $|F(calc)|$ are the observed and calculated structure factor amplitudes, respectively.

Parameter	GP*a*	
	Activated	Inhibited
	Crystal data	
Unit cell dimensions		
a (Å)	128.5	128.4
c (Å)	118.1	116.4
Resolution (Å)	2.8	2.1
Reflections	23,035	35,231
R_{sym}	0.06	0.05
	Refinement parameters	
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
R_{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.