Vol. 166, No. 1, 1990 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
January 15, 1990 Pages 394-403

CLONING, SEQUENCING AND EXPRESSION OF HUMAN TSH RECEPTOR

Micheline MISRAHI, Hugues LOOSFELT, Michel ATGER,
Sokhavuth SAR, Anne GUIOCHON-MANTEL and Edwin MILGROM*

Institut National de la Santé et de la Recherche Médicale Unité 135,
Hormones et Reproduction, Hôpital de Bicêtre,
94270 LE KREMLIN BICETRE, FRANCE

Received December 9, 1989

Complementary cDNA clones encoding the TSH (thyroid stimulatory hormone) receptor were isolated from a human thyroid λgt10 library using low stringency hybridization with LH/hCG (luteinizing hormone-human choriogonadotropic hormone) receptor probes. Sequencing of the clones showed a 764 amino acid open reading frame. The first 21 amino acids probably correspond to a signal peptide, the mature protein thus contains 743 amino acids (calculated molecular weight: 84,501 daltons). Its putative structure consists of a 394 amino acid extracellular domain, a 266 amino acid membrane spanning domain with 7 putative transmembrane segments and a 83 amino acid intracellular domain. A high degree of homology is observed with LH/hCG receptor suggesting the definition of a new subfamily of G-protein coupled receptors. Computer search showed the presence in the putative third intracellular loop of a motif resembling that described in the non receptor type protein tyrosine kinases (c-src, c-yes, c-fgr, etc...). RNA blots showed that the receptor messenger RNA consists of two major species of 4300 and 3900 nucleotides. The cDNA was inserted into an expression vector and after transfection into COS 7 cells it was shown to produce a functional TSH receptor. © 1990 Academic Press, Inc.

The TSH (thyroid stimulatory hormone) receptor has been the subject of numerous studies (reviewed in 1) but its structure is still not understood: molecular weights varying between 30,000 and 200,000 have been reported. Several studies have suggested that the receptor is an hetero-oligomer containing different subunits of which only one binds the hormone. The interest in TSH receptor has also come from clinical considerations. Graves' disease is a very frequent ailment, especially in women (2), and is caused by the occurence of stimulatory anti-TSH receptor antibodies (review in 3). The reason for the appearance of these antibodies is not understood: genetic predisposition (predominance of certain HLA haplotypes at the B and D loci) and aberrantly regulated

* To whom correspondence should be addressed.

0006-291X/90 $1.50

idiotypic network have been implicated. Moreover, blocking antireceptor antibodies have been demonstrated in idiopathic myxoedema (4) and in endemic cretinism (5). Thus progress in the understanding of thyroid physiology and pathogenesis as well as improvement in diagnosis and management of thyroid disease requires the isolation and the characterization of human TSH receptor.

METHODS cDNA cloning: Random primed cDNAs were prepared from human thyroid polyadenylated RNAs. Size selection of the cDNAs and cloning into λgt10 vector were as described (6). The cDNA library (1.5×10^6 clones) was screened with nick translated ^{32}P labeled full length porcine LH/hCG receptor cDNA (6). Nitrocellulose filters were hybridized at 37° for 36 hours with 5.10^5 cpm/ml of probe in 6xSSC (1xSSC = 0.15 M Sodium chloride, 0.015 M Sodium Citrate), buffer containing 30 % formamide, 0.5 % polyvinylpyrrolidone, 0.5 % Ficoll, 0.5 % bovine serum albumin, denatured Salmon sperm DNA (150 µg/ml) and 0.2 % sodium dodecyl sulfate. Filters were washed twice for 15 min at 50°C in 6xSSC and twice for 15 min at 25°C in 2xSSC and finally 10 minutes at 45°C in 2xSSC. One hundred positive clones were examined in further experiments for their capacity to hybridize with either the extracellular or the transmembrane encoding region of the porcine LH/hCG receptor cDNA. Eight clones were selected for sequence analysis.

Sequencing was performed as described (7) after subcloning λTSHR clones into Blue Script vector (Stratagene).

RNA blot analyses were performed using polyadenylated RNAs from human tisues. Hybridization was performed with a random primed 2.2 Kb cDNA insert of $\lambda TSHR_3$ (6×10^6 cpm/ng, 5×10^5 cpm/ml).

Expression of the cloned cDNA: A full length coding region for hTSHR was constructed using $\lambda hTSHR_3$ and $\lambda hTSHR_4$ clones (it extends from nucleotide - 44 to nucleotide + 2395, the numbering starting at the first codon). It was inserted into the pKSV10 vector (Pharmacia) and used to transfect Cos-7 cells as described (8). Cells were used 48 hours after transfection.

^{125}I-bTSH binding studies: Purified bovine TSH (40 i.u/mg), a gift from Dr John Pierce (University of California, Los Angeles), was iodinated by the lactoperoxydase method (9) and was purified by Ultrogel ACA44 chromatography. Specific activity was 25 µCi/µg and maximal binding activity was 30% of total radioactivity (measured on porcine thyroid membranes). Binding assays were performed in triplicate with 10^5 cells in 200 µl of PSA (20 mM phosphate pH 7.4, 0.2% bovine serum albumin) 0.25 M sucrose buffer containing 3×10^4 cpm ^{125}I-bTSH and various amounts of bTSH, (30 i.u./mg) (gift of the National Hormone and Pituitary program, NHPP, Baltimore). After 30 min incubation at 37°C the suspension was layered onto 400 µl of PSA 1 M sucrose buffer and centrifuged at 10,000 g for 10 min. The supernatants were aspirated and the bottoms of the tubes were cut out and counted for radioactivity. Control experiments were performed with cells transfected with rabbit progesterone receptor cDNA expression vector (8). To test the specificity of the bTSH binding, competition experiments were also performed using unlabeled purified porcine LH (YC 1781, gift from Yves Combarnous, Nouzilly, France) and human FSH (150 i.u./mg, Metrodin Serono - Switzerland).

Stimulation of adenylate cyclase activity by TSH: 6×10^5 COS 7 cells were plated in 60 nM dishes and transfected with 5 µg of pKSV-TSHR and 5 µg of carrier DNA. 42 hours later, each dish was washed three times with warm Eagle medium containing gelatine (1 mg/ml). Each dish was then incubated with 4.5 ml of the same medium containing 0.5 nM 3-isobutyl-1-methyl xanthine during 15 minutes at 37° C. Various concentrations of bTSH were added and the incubation was continued for 5 minutes at 37° C. The medium was removed and the cells were collected into 0.8 ml of 1 N perchloric

acid and centrifuged. The supernantants were neutralised and assayed for cAMP by radioimmunoassay (International CIS).

RESULTS AND DISCUSSION Cloning of TSH receptor cDNA: Poly (A+) RNAs were isolated from human thyroids and used to prepare a cDNA library in λgt10 vector. The library was screened at low stringency, with a probe corresponding to porcine LH/hCG receptor (6). The rationale for this approach was based on the similarity in the structure of the ligands (LH and TSH) (10), suggesting that their receptors could belong to a family of cross-hybridizing genes. Indeed screening 1.5×10^6 clones of the library led to the isolation of 180 positive signals. Eight clones were selected for sequence analysis (Fig 1C) using probes corresponding to 5' and 3' parts of the porcine LH/hCG receptor cDNA.

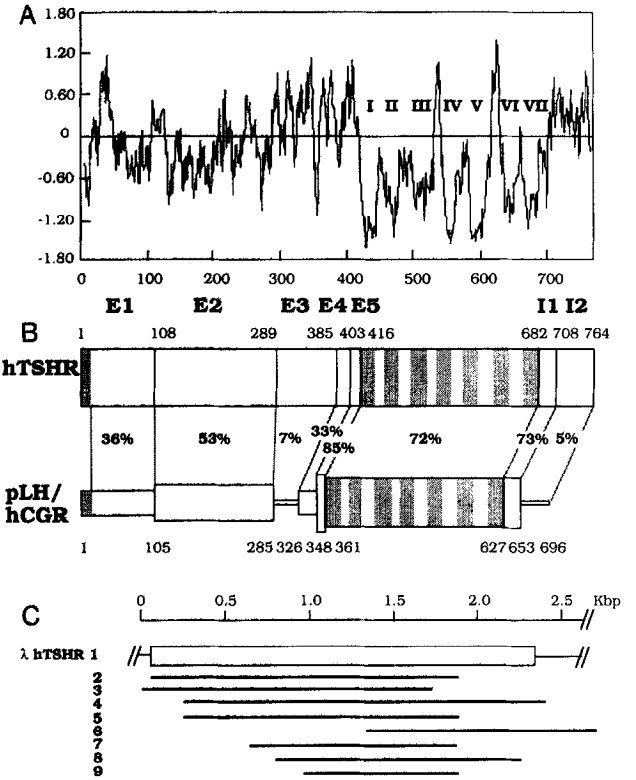


Fig. 1. Schematic representation of the structure of human TSH receptor. A. Hydropathy plot (23). Hydrophobic regions correspond to negative values. I to VII correspond to the putative transmembrane segments. B. Comparison between hTSH and pLH/hCG receptors (6). Receptors are divided into regions according to the extent of homology (marked by the thickness of the figure representing the LH/hCG receptor). Dashed regions represent the putative signal peptide and the seven membrane spans. E1-5 are the putative extracellular and I1,2 intracellular domains. Amino acids numbering is shown above and below the figure. C. Map of cDNA clones used for the sequencing. λhTSHR are the λgt10 clones. The open reading frame is boxed. Kbp=Kilobase pair.

hTSH R

```
                                                                                                                    ggcgatttcggaggatggagaaatagccccgagtcccgtggaaa   -1
                                                                                C
ATG AGG CCG GCG GAC TTG CTG CAG CTG GTG CTG CTG CTC GAC CTG CCC AGG GAC CTG GGC GGA ATG GGG TGT TCG TCT CCA CCC    84
Met Arg Pro Ala Asp Leu Leu Gln Leu Val Leu Leu Leu Asp Leu Pro Arg Asp Leu Gly Gly Met Gly Cys Ser Ser Pro Pro    28

TGC GAG TGC CAT CAG GAG GAG GAC TTC AGA GTC ACC TGC AAG GAT ATT CAA CGC ATC CCC AGC TTA CCG CCC AGT ACC CAG ACT   168
Cys Glu Cys His Gln Glu Glu Asp Phe Arg Val Thr Cys Lys Asp Ile Gln Arg Ile Pro Ser Leu Pro Pro Ser Thr Gln Thr    56

CTG AAG CTT ATT GAG ACT CAC CTG AGA ACT ATT CCA AGT CAT GCA TTT TCT AAT CTG CCC AAT ATT TCC AGA ATC TAC GTA TCT   252
Leu Lys Leu Ile Glu Thr His Leu Arg Thr Ile Pro Ser His Ala Phe Ser Asn Leu Pro Asn Ile Ser Arg Ile Tyr Val Ser    84

ATA GAT GTG ACT CTG CAG CAG CTG GAA TCA CAC TCC TTC TAC AAT TTG AGT AAA GTG ACT CAC ATA GAA ATT CGG AAT ACC AGG   336
Ile Asp Val Thr Leu Gln Gln Leu Glu Ser His Ser Phe Tyr Asn Leu Ser Lys Val Thr His Ile Glu Ile Arg Asn Thr Arg   112

AAC TTA ACT TAC ATA GAC CCT GAT GCC CTC AAA GAG CTC CCC CTC CTA AAG TTC CTT GGC ATT TTC AAC ACT GGA CTT AAA ATG   420
Asn Leu Thr Tyr Ile Asp Pro Asp Ala Leu Lys Glu Leu Pro Leu Leu Lys Phe Leu Gly Ile Phe Asn Thr Gly Leu Lys Met   140

TTC CCT GAC CTG ACC AAA GTT TAT TCC ACT GAT ATA TTC TTT ATA CTT GAA ATT ACA GAC AAC CCT TAC ATG ACG TCA ATC CCT   504
Phe Pro Asp Leu Thr Lys Val Tyr Ser Thr Asp Ile Phe Phe Ile Leu Glu Ile Thr Asp Asn Pro Tyr Met Thr Ser Ile Pro   168

GTG AAT GCT TTT CAG GGA CTA TGC AAT GAA ACC TTG ACA CTG AAG CTG TAC AAC AAC GGC TTT ACT TCA GTC CAA GGA TAT GCT   588
Val Asn Ala Phe Gln Gly Leu Cys Asn Glu Thr Leu Thr Leu Lys Leu Tyr Asn Asn Gly Phe Thr Ser Val Gln Gly Tyr Ala   196

TTC AAT GGG ACA AAG CTG GAT GCT GTT TAC CTA AAC AAG AAT AAA TAC CTG ACA GTT ATT GAC AAA GAT GCA TTT GGA GGA GTA   672
Phe Asn Gly Thr Lys Leu Asp Ala Val Tyr Leu Asn Lys Asn Lys Tyr Leu Thr Val Ile Asp Lys Asp Ala Phe Gly Gly Val   224

TAC AGT GGA CCA AGC TTG CTG GAC GTG TCT CAA ACC AGT GTC ACT GCC CTT CCA TCC AAA GGC CTG GAG CAC CTG AAG GAA CTG   756
Tyr Ser Gly Pro Ser Leu Leu Asp Val Ser Gln Thr Ser Val Thr Ala Leu Pro Ser Lys Gly Leu Glu His Leu Lys Glu Leu   252
                                C                                                                   C
ATA GCA AGA AAC ACC TGG ACT CTT AAG AAA CTT CCA CTT TCC TTG AGT TTC CTT CAC CTC ACA CGG GCT GAC CTT TCT TAC CCA   840
Ile Ala Arg Asn Thr Trp Thr Leu Lys Lys Leu Pro Leu Ser Leu Ser Phe Leu His Leu Thr Arg Ala Asp Leu Ser Tyr Pro   280
                    T
AGC CAC TGC TGT GCC TTT AAG AAT CAG AAG AAA ATC AGA GGA ATC CTT GAG TCC TTG ATG TGT AAT GAG AGC AGT ATG CAG AGC   924
Ser His Cys Cys Ala Phe Lys Asn Gln Lys Lys Ile Arg Gly Ile Leu Glu Ser Leu Met Cys Asn Glu Ser Ser Met Gln Ser   308

TTG CGC CAG AGA AAA TCT GTG AAT GCC TTG AAT AGC CCC CTC CAC CAG GAA TA1 GAA GAG AAT CTG GGT GAC AGC ATT GTT GGG  1008
Leu Arg Gln Arg Lys Ser Val Asn Ala Leu Asn Ser Pro Leu His Gln Glu Tyr Glu Glu Asn Leu Gly Asp Ser Ile Val Gly   336

TAC AAG GAA AAG TCC AAG TTC CAG GAT ACT CAT AAC AAC GCT CAT TAT TAC GTC TTC TTT GAA GAA CAA GAG GAT GAG ATC ATT  1092
Tyr Lys Glu Lys Ser Lys Phe Gln Asp Thr His Asn Asn Ala His Tyr Tyr Val Phe Phe Glu Glu Gln Glu Asp Glu Ile Ile   364

GGT TTT GGC CAG GAG CTC AAA AAC CCC CAG GAA GAG ACT CTA CAA GCT TTT GAC AGC CAT TAT GAC TAC ACC ATA TGT GGG GAC  1176
Gly Phe Gly Gln Glu Leu Lys Asn Pro Gln Glu Glu Thr Leu Gln Ala Phe Asp Ser His Tyr Asp Tyr Thr Ile Cys Gly Asp   392
                                                                                    C
AGT GAA GAC ATG GTG TGT ACC CCC AAG TCC GAT GAG TTC AAC CCG TGT GAA GAC ATA ATG GGC TAC AAG TTC CTG AGA ATT GTG  1260
Ser Glu Asp Met Val Cys Thr Pro Lys Ser Asp Glu Phe Asn Pro Cys Glu Asp Ile Met Gly Tyr Lys Phe Leu Arg Ile Val   420

GTG TGG TTC GTT AGT CTG CTG GCT CTC CTG GGC AAT GTC TTT GTC CTG CTT ATT CTC CTC ACC AGC CAC TAC AAA CTG AAC GTC  1344
Val Trp Phe Val Ser Leu Leu Ala Leu Leu Gly Asn Val Phe Val Leu Leu Ile Leu Leu Thr Ser His Tyr Lys Leu Asn Val   448

CCC CGC TTT CTC ATG TGC AAC CTG GCC TTT GCG GAT TTC TGC ATG GGG ATG TAC CTG CTC CTC ATC GCC TCT GTA GAC CTC TAC  1428
Pro Arg Phe Leu Met Cys Asn Leu Ala Phe Ala Asp Phe Cys Met Gly Met Tyr Leu Leu Leu Ile Ala Ser Val Asp Leu Tyr   476
                                                                                                C
ACT CAC TCT GAG TAC TAC AAC CAT GCC ATC GAC TGG CAG ACA GGC CCT GGG TGC AAC ACG GCT GGT TTC TTC ACT GTC TTT GCA  1512
Thr His Ser Glu Tyr Tyr Asn His Ala Ile Asp Trp Gln Thr Gly Pro Gly Cys Asn Thr Ala Gly Phe Phe Thr Val Phe Ala   504

AGC GAG TTA TCG GTG TAT ACG CTG ACG GTC ATC ACC CTG GAG CGC TGG TAT GCC ATC ACC TTC GCC ATG CGC CTG GAC CGG AAG  1596
Ser Glu Leu Ser Val Tyr Thr Leu Thr Val Ile Thr Leu Glu Arg Trp Tyr Ala Ile Thr Phe Ala Met Arg Leu Asp Arg Lys   532

ATC CGC CTC AGG CAC GCA TGT GCC ATC ATG GTT GGG GGC TGG GTT TGC TGC TTC CTT CTC GCC CTG CTT CCT TTG GTG GGA ATA  1680
Ile Arg Leu Arg His Ala Cys Ala Ile Met Val Gly Gly Trp Val Cys Cys Phe Leu Leu Ala Leu Leu Pro Leu Val Gly Ile   560

AGT AGC TAT GCC AAA GTC AGT ATC TGC CTG CCC ATG GAC ACC GAG ACC CCT CTT GCT CTG GCA TAT ATT GTT TTT GTT CTG ACG  1764
Ser Ser Tyr Ala Lys Val Ser Ile Cys Leu Pro Met Asp Thr Glu Thr Pro Leu Ala Leu Ala Tyr Ile Val Phe Val Leu Thr   588
                                                C
CTC AAC ATA GTT GCC TTC GTC ATC GTC TGC TGC TGT TAT GTG AAG ATC TAC ATC ACA GTC CGA AAT CCG CAG TAC AAC CCA GGG  1848
Leu Asn Ile Val Ala Phe Val Ile Val Cys Cys Cys Tyr Val Lys Ile Tyr Ile Thr Val Arg Asn Pro Gln Tyr Asn Pro Gly   616
                                                            C
GAC AAA GAT ACC AAA ATT GCC AAG AGG ATG GCT GTG TTG ATC TTC ACC GAC TTC ATA TGC ATG GCC CCA ATC TCA TTC TAT GCT  1932
Asp Lys Asp Thr Lys Ile Ala Lys Arg Met Ala Val Leu Ile Phe Thr Asp Phe Ile Cys Met Ala Pro Ile Ser Phe Tyr Ala   644

CTG TCA GCA ATT CTG AAC AAG CCT CTC ATC ACT GTT AGC AAC TCC AAA ATC TTG CTG GTA CTC TTC TAT CCA CTT AAC TCC TGT  2016
Leu Ser Ala Ile Leu Asn Lys Pro Leu Ile Thr Val Ser Asn Ser Lys Ile Leu Leu Val Leu Phe Tyr Pro Leu Asn Ser Cys   672

GCC AAT CCA TTC CTC TAT GCT ATT TTC ACC AAG GCC TTC CAG AGG GAT GTG TTC ATC CTA CTC AGC AAG TTT GGC ATC TGT AAA  2100
Ala Asn Pro Phe Leu Tyr Ala Ile Phe Thr Lys Ala Phe Gln Arg Asp Val Phe Ile Leu Leu Ser Lys Phe Gly Ile Cys Lys   700
                                                                                                            C
CGC CAG GCT CAG GCA TAC CGG GGG CAG AGG GTT CCT CCA AAG AAC AGC ACT GAT ATT CAG GTT CAA AAG GTT ACC CAC GAG ATG  2184
Arg Gln Ala Gln Ala Tyr Arg Gly Gln Arg Val Pro Pro Lys Asn Ser Thr Asp Ile Gln Val Gln Lys Val Thr His Glu Met   728
                                                        C
AGG CAG GGT CTC CAC AAC ATG GAA GAT GTC TAT GAA CTG ATT GAA AAG TCC CAT CTA ACC CCA AAG AAG CAA GGC CAA ATC TCA  2268
Arg Gln Gly Leu His Asn Met Glu Asp Val Tyr Glu Leu Ile Glu Lys Ser His Leu Thr Pro Lys Lys Gln Gly Gln Ile Ser   756
                                                                ----------
GAA GAG TAT ATG CAA ACG GTT TTG  taagttaacactacactactcacaatggtaggggaacttacaaaatatagtttcttgaatatgcattccaatcccatg 2371
Glu Glu Tyr Met Gln Thr Val Leu                                                                                 764
```

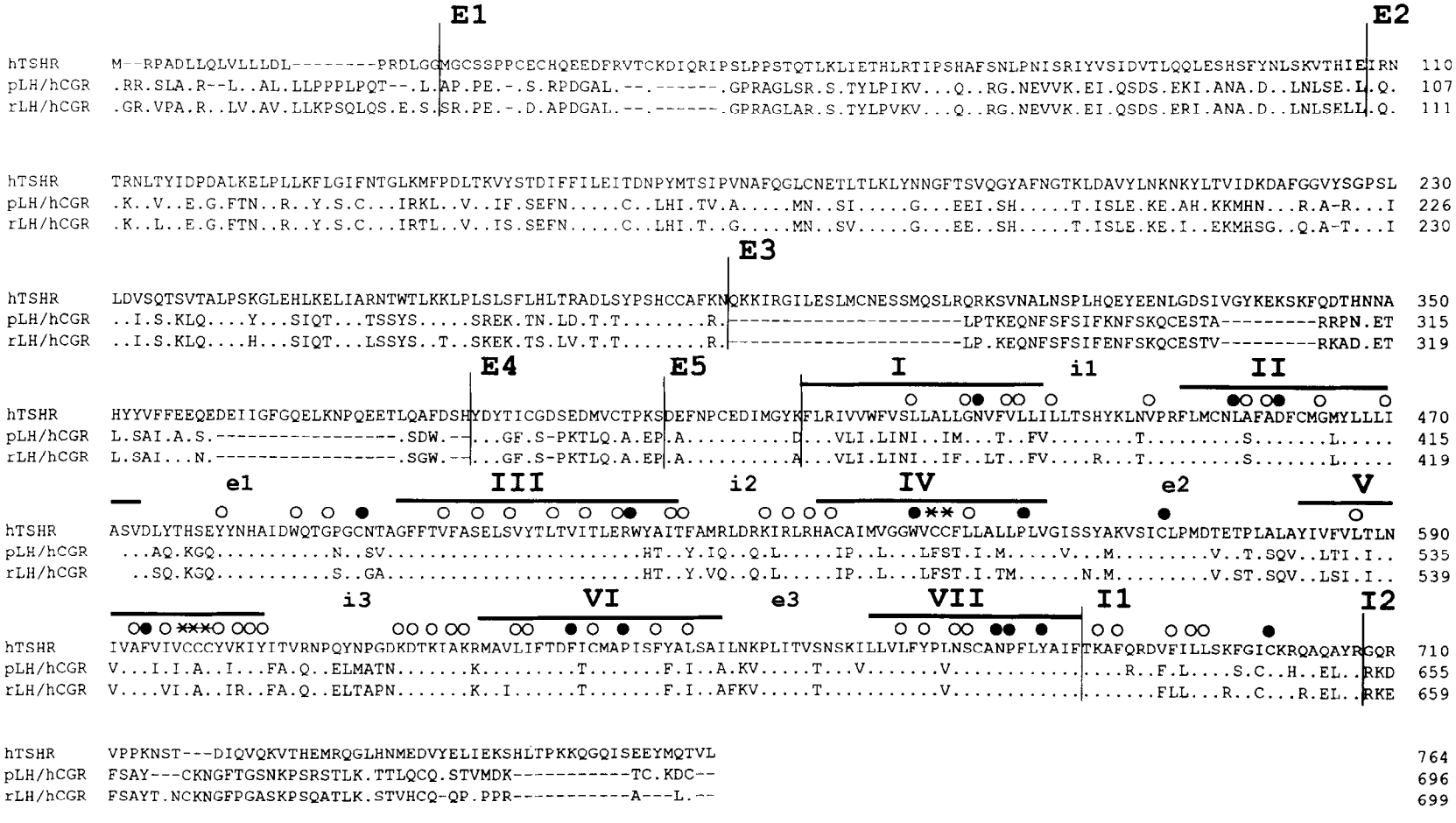
Fig. 2. Sequence of the human TSH receptor (hTSHR) and of the deduced protein: Position + 1 is assigned to the first nucleotide of the putative initiator codon. Numbering of the nucleotides (above) and of the aminoacids (underneath) is shown. The potential N-linked glycosylation sites in the extracellular domain are underlined, a putative site for kinase C phosphorylation is indicated by a dotted line. Ponctual divergences from the main sequence were found in individual clones. They are indicated above the sequence. Changes in the encoded amino acids are Leu ---> Pro 13, Leu ---> Pro 260, Tyr ---> His 414, Phe ---> Leu 500, Phe ---> Leu 634, Glu ---> Asp 727, Lys ---> Asn 744.

Sequence analysis of TSH receptor: The nucleotide sequence shows an ATG which is preceeded by an upstream inframe stop codon thus defining an open reading frame of 764 amino acids (Fig 2). The N-terminal end encodes a 21 amino acid sequence characteristic of a signal peptide with a cleavage site as defined by Von Hejne (11). Thus the mature protein probably consists of 743 amino acids with a calculated molecular weight of 84,501 daltons. Comparison with other G-protein linked receptors (12,13,14,15) including the LH/hCG receptor (6,16) and the hydropathy profile of TSH receptor (Fig 1A) suggest the following probable structural organization of the protein: a large putative extracellular domain extends over 394 amino acids of the N-terminal part. It contains 6 putative N-linked glycosylation sites. Comparison with the LH/hCG receptor suggests that this region is divided into five segments (Fig 1B) of which one (E5) exhibits a high homology (85 %) and three others a somewhat lower homology. A fifth domain (E3) is highly acidic (pKi=4.14) and diverges by sequence and length from the corresponding and equally acidic shorter region of the LH receptor. Pituitary glycoproteins (LH, TSH, FSH) share a commun α subunit and have related but specific β subunits. Since it has been proposed that the extracellular domain of LH receptor is involved in ligand binding (6,16) it is tempting to speculate that the most conserved regions interact with the α subunit and the most divergent regions with the β subunit. Two clusters of cysteins are found on both extremities of the putative extracellular domain. They are completely conserved between LH and TSH receptors (Fig 3).

The putative membrane spanning domain is 266 amino acids long and contains the characteristic pattern of seven probable transmembrane segments. It exhibits a high overall homology (72 %) with the LH/hCG receptor and a lower but significant homology with the corresponding regions of other G-protein linked receptors (Fig 3). The homology with the LH/hCG receptor is especially focused on the II, III, VI and VII putative

---

Fig. 3. Detailed comparison between the TSH and LH/hCG receptors: Only non identical amino acids of porcine (6) and rat (16) LH/hCG receptors are shown, identical residues are represented by a dot. Alignment of homologous regions has necessitated introduction of gaps represented by dashes. The numbering of the amino acids is indicated on the right. The limits of the domains described in Figure 1 are indicated. The putative transmembrane segments I to VII are shown by a thick line above the sequence. e1 to3 are the putative extracellular and i 1 to3 intracellular loops. The group of cysteins characteristic of transmembrane domains IV and V are shown by stars. The amino acids conserved in more than half of the various G-protein coupled receptors (bovine Rhodopsin (12), human β1 (13), β2 (14), β3 (15) hamster α1 (24) and human α2 (25) adrenergic, human M1-M4 (26) and rat M5 muscarinic (27), rat dopaminergic (28), human 5HT1A (29) rat 5HT1C (30) and 5HT2 (31) and bovine substance K (32) receptors) are indicated by a an open circle (K and R, D and E, I and L are considered equivalent). The amino acids which are conserved in all receptors are indicated by a full circle.


E1
E2
E3
E4
E5
I
i1
II
e1
III
i2
IV
e2
V
i3
VI
e3
VII
I1
I2
hTSHR M--RPADLLQLVLLLDL--------PRDLGCMGCSSPPCECHQEEDFRVTCKDIQRIPSLPPSTQTLKLIETHLRTIPSHAFSNLPNISRIYVSIDVTLQQLESHSFYNLSKVTHIEIRN 110
pLH/hCGR .RR.SLA.R--L..AL.LLPPPLPQT--.L.AP.PE.-.S.RPDGAL.--.-------.GPRAGLSR.S.TYLPIKV...Q..RG.NEVVK.EI.QSDS.EKI.ANA.D..LNLSE.L.Q. 107
rLH/hCGR .GR.VPA.R..LV.AV.LLKPSQLQS.E.S.SR.PE.-.D.APDGAL.--.-------.GPRAGLAR.S.TYLPVKV...Q..RG.NEVVK.EI.QSDS.ERI.ANA.D..LNLSELL.Q. 111
hTSHR TRNLTYIDPDALKELPLLKFLGIFNTGLKMFPDLTKVYSTDIFFILEITDNPYMTSIPVNAFQGLCNETLTLKLYNNGFTSVQGYAFNGTKLDAVYLNKNKYLTVIDKDAFGGVYSGPSL 230
pLH/hCGR .K..V..E.G.FTN..R..Y.S.C...IRKL..V..IF.SEFN.....C..LHI.TV.A.....MN..SI.....G...EEI.SH.....T.ISLE.KE.AH.KKMHN...R.A-R...I 226
rLH/hCGR .K..L..E.G.FTN..R..Y.S.C...IRTL..V..IS.SEFN.....C..LHI.T..G.....MN..SV.....G...EE..SH.....T.ISLE.KE.I..EKMHSG..Q.A-T...I 230
hTSHR LDVSQTSVTALPSKGLEHLKELIARNTWTLKKLPLSLSFLHLTRADLSYPSHCCAFKNQKKIRGILESLMCNESSMQSLRQRKSVNALNSPLHQEYEENLGDSIVGYKEKSKFQDTHNNA 350
pLH/hCGR ..I.S.KLQ....Y...SIQT...TSSYS.....SREK.TN.LD.T.T.....R.----------LPTKEQNFSFSIFKNFSKQCESTA---------RRPN.ET 315
rLH/hCGR ..I.S.KLQ....H...SIQT...LSSYS..T..SKEK.TS.LV.T.T.....R.----------LP.KEQNFSFSIFENFSKQCESTV---------RKAD.ET 319
hTSHR HYYVFFEEQEDEIIGFGQELKNPQEETLQAFDSHYDYTICGDSEDMVCTPKSDEFNPCEDIMGYKFLRIVVWFVSLLALLGNVFVLLILLTSHYKLNVPRFLMCNLAFADFCMGMYLLLI 470
pLH/hCGR L.SAI.A.S.----------.SDW.--...GF.S-PKTLQ.A.EP.A.....D...VLI.LINI..IM...T..FV.....T.....S.....L..... 415
rLH/hCGR L.SAI...N.----------.SGW.--...GF.S-PKTLQ.A.EP.A.....A...VLI.LINI..IF..LT..FV....R...T.....S.....L..... 419
hTSHR ASVDLYTHSEYYNHAIDWQTGPGCNTAGFFTVFASELSVYTLTVITLERWYAITFAMRLDRKIRLRHACAIMVGGWVCCFLLALLPLVGISSYAKVSICLPMDTETPLALAYIVFVLTLN 590
pLH/hCGR ...AQ.KGQ.....N..SV.....HT..Y.IQ..Q.L.....IP..L...LFST.I.M.....V...M.....V..T.SQV..LTI.I.. 535
rLH/hCGR ..SQ.KGQ.....S..GA.....HT..Y.VQ..Q.L.....IP..L...LFST.I.TM.....N.M.....V.ST.SQV..LSI.I.. 539
hTSHR IVAFVIVCCCYVKIYITVRNPQYNPGDKDTKIAKRMAVLIFTDFICMAPISFYALSAILNKPLITVSNSKILLVLFYPLNSCANPFLYAIFTKAFQRDVFILLSKFGICKRQAQAYRGQR 710
pLH/hCGR V...I.I.A..I...FA.Q..ELMATN.....K.....T.....F.I..A.KV.....T...V.....V.....|....R..F.L....S.C..H..EL..RKD 655
rLH/hCGR V....VI.A..IR..FA.Q..ELTAPN.....K..I.....T.....F.I..AFKV.....T.....V.....|.....FLL...R..C...R.EL..RKE 659
hTSHR VPPKNST---DIQVQKVTHEMRQGLHNMEDVYELIEKSHLTPKKQGQISEEYMQTVL 764
pLH/hCGR FSAY---CKNGFTGSNKPSRSTLK.TTLQCQ.STVMDK----------TC.KDC-- 696
rLH/hCGR FSAYT.NCKNGFPGASKPSQATLK.STVHCQ-QP.PPR----------A---L.-- 699

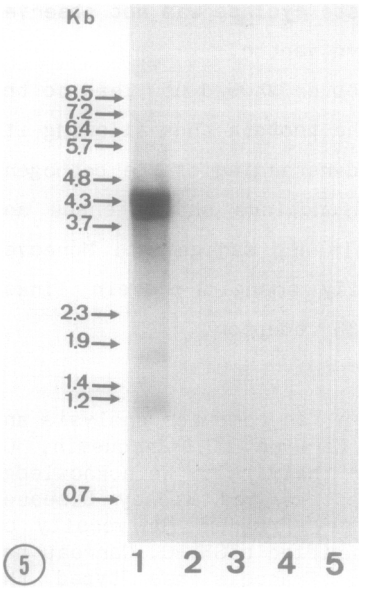
transmembrane segments (Fig 3). This pattern is similar to that observed when comparing subtypes of β adrenergic receptors (15). The putative extracellular loops are well conserved between LH/hCG and TSH receptors specially in their central parts. The first putative intracellular loop is highly conserved. On the contrary the amino terminal part of the third intracellular loop is specific to the TSH receptor. Remarkably a computer search showed this region of the TSH receptor to share 8 amino acids with the carboxy-terminal twelve aminoacid motif found in all the protein tyrosine kinases of the non-receptor type (c-src (17), c-yes, (18) c-fgr (19) etc...). A tyrosine within this motif has been shown to be phosphorylated in c-src and to play a role in regulating pp60 c-src kinase activity (Fig.4). This short sequence is absent in the corresponding oncogenic retroviruses. This alteration has been suggested to enhance src transforming activity (20). In adrenergic receptors the third loop together with transmembrane domains V and VI have been implicated in specific interactions with G-proteins (review in 21).

The putative intracellular part of the TSH receptor is 83 amino acids long and very basic (pKi= 9.6). A first segment is highly conserved between LH/hCG and TSH receptors (Fig.1B,3). A second segment is divergent. However, its N-terminal part is highly conserved when comparing LH/hCG receptors from pig and rat, this suggests the presence of a hormone specific function (Fig.3). On the contrary its C terminal part is variable in all receptors. A high proportion of serines and threonines are found in the putative intracellular domain with a consensus site for phosphorylation by protein kinase C (Fig.2). Phosphorylation by specific kinases plays a role in the agonist specific decoupling of adrenergic receptors from G-proteins (22).

Some clones diverged from the main sequence. Further analysis will be necessary to establish if they correspond to variant forms of receptor as observed for the porcine LH/hCG receptor (6) or to cloning artefacts. The cloning strategy which involved hybridization with LH/hCG receptor probes may have prevented the isolation of putative clones lacking the most conserved transmembrane domain.

hTSHR	Y	I	T	V	R	N	P	Q	Y	N	P	G	D	K	D			605-619
c-fgr	Y	F	T	S	A	E	P	Q	Y	Q	P	G	D	Q	T	-	COOH	515-529
c-slk	Y	F	T	A	T	E	P	Q	Y	Q	P	G	E	N	L	-	COOH	523-537
c-syn	Y	F	T	A	T	E	P	Q	Y	Q	P	G	E	N	L	-	COOH	523-537
c-yes	Y	F	T	A	T	E	P	Q	Y	Q	P	G	E	N	L	-	COOH	529-543
c-src	Y	F	T	S	T	E	P	Q	Y	Q	P	G	E	N	L	-	COOH	519-543
v-src	A	C	V	L	E	V	A	E	-	COOH								519-526

Fig. 4. Similarity between a domain of human TSH receptor and the C-terminal part of non-receptor tyrosine protein kinases. Conserved aminoacids are boxed.



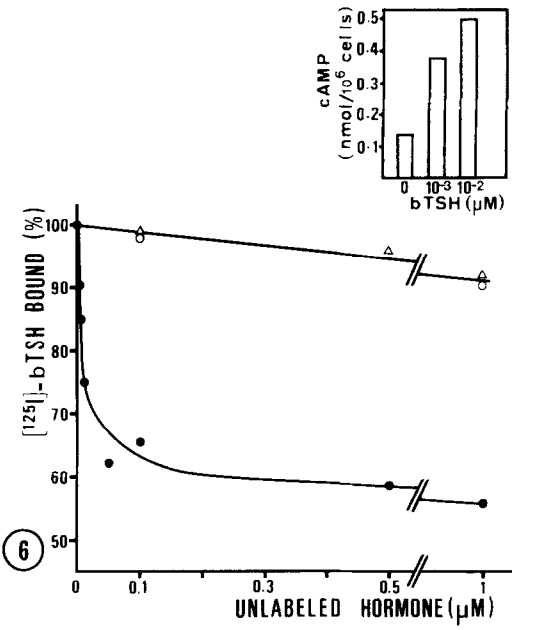


Fig. 5. RNA blot analysis of TSH receptor messenger RNA. Polyadenylated RNAs (20 ug) from human thyroid (lane 1), testes (lane 2), ovaries (lane 3), liver (lane 4) and spleen (lane 5) were analyzed. Size of DNA markers are indicated (Kb=kilobase).

Fig. 6. Expression of TSH receptor cDNA in COS 7 cells. COS 7 cells were transfected with a vector encoding TSH receptor. Binding of ^{125}I-bTSH was studied as described in Methods in presence of increasing concentrations of competing unlabeled bTSH (●), pLH (○) and hFSH (△). Results are shown as percent of ^{125}I-bTSH bound in the absence of competing unlabeled hormone (means of triplicate experiments). Inset: Stimulation of adenylate cyclase (see Methods).

Northern blot analysis of poly (A+) RNAs (Fig 5) and total RNAs (not shown) from several organs showed the presence in the thyroid of a major messenger species of 4300 nucleotides and of a less abundant band of 3900 nucleotides. Smaller and minor species of 1700 and 1100 nucleotides could also be seen. As expected no messenger was observed in the testes, ovaries, spleen and liver.

Expression of the cloned cDNA: A full length coding region was reconstructed from λhTSHR$_2$ and λhTSHR$_3$ (Fig 1C), inserted into the pKSV10 expression vector and transfected into COS 7 cells. This led to the appearance on the membranes of cells of a protein which bound ^{125}I-bTSH. This binding was saturable (displacement by unlabeled TSH) and specific (^{125}I-bTSH could not be displaced by unlabeled pLH or hFSH) (Fig.6). Moreover no saturable binding of ^{125}I-bTSH could be observed when cells were transfected with a control pKSV10 expression vector encoding rabbit progesterone receptor (data not shown). Incubation with TSH of the cells transfected with the pKSV-TSHR vector led to increased accumulation of

cAMP (Fig.6 Inset). Such stimulation of adenylate cyclase was not observed in cells transfected with the control expression vector.

The isolation of the cDNA encoding TSH receptor should now lead to the preparation of monoclonal antibodies against the protein thus allowing its detailed characterization and purification. Understanding of the pathogeny of Graves' disease and of other thyroid dysfunctions should ensue and progress should be made in methods of diagnosis and management. Moreover LH/hCG and TSH receptors define a new family among G-protein linked receptors to which probably also belongs the FSH receptor.

ACKNOWLEDGMENTS We are grateful to Dr P. Dessen for computer analysis and helpful discussions. We also thank Drs P. Carayon, C. Jacquemin, C. Correze, B. Haye, J. Reynier and D. Chopin for their help. We acknowledge the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) and the National Hormone and Pituitary Program (University of Maryland School of Medicine) for the gift of purified bTSH. C. Carreau has provided excellent technical assistance. V. Francois has typed the manuscript. This work was supported by the INSERM, the University Paris-South, the Fondation pour la Recherche Medicale Française and Transbio Company.

REFERENCES

1. Farid, N.R., and Fahraeus-Van Ree, G. (1986) Immunoassay Technology, Pal, S. (ed.) pp. 195-220.
2. Farid, N.R. (1987) Endocrinol. Metab. Clin. North. Am. 16, 229-245.
3. Rees Smith, B., McLachlan, S.M., and Furmaniak, J. (1988) Endocr. Rev.9, 106-121.
4. Matsuura, N., Yamada, Y., Nohara, Y., Konishi, J., Kasagi, K., Endo, K.,Kojima, H., and Wataya, K. (1980) New Engl. J. Med. 303, 738-741.
5. Boyages, S.C., Halpern, J.P., Maberly, G.F., Eastman, G.J., Chen, J., Zhen-Hua, W., VanDer Gaag, R.D., and Drexhage, H.A. (1989) Lancet 529-532.
6. Loosfelt, H., Misrahi, M., Atger, M., Salesse, R., Vu Hai, M.T., Jolivet, A., Guiochon-Mantel, A., Sar, S., Jallal, B., Garnier, J., and Milgrom, E. (1989) Science 245, 525-528.
7. Misrahi, M., Atger, M., d'Auriol, L., Loosfelt, H., Meriel, C., Fridlansky, F., Guiochon-Mantel, A., Galibert, F., and Milgrom E. (1987) Biochem. Biophys. Res. Commun. 143, 740-748.
8. Guiochon-Mantel, A., Loosfelt, H., Ragot, T., Bailly, A., Atger, M., Misrahi, M., Perricaudet, M., and Milgrom, E. (1988) Nature 336, 695-698.
9. Powell-Jones, C.H.J., Thomas, Jr.C.G., and Nayfeh, S.N. (1980) J. Biol. Chem. 255, 4001-4010.
10. Kourides, I.A., Gurr, J.A., and Wolf, O. (1984) Recent Progres in Hormone Research, Academic Press 40, pp. 79-117.
11. Von Heijne, G. (1986) Nucleic Acids Res. 14, 4683-4686.
12. Nathans, J., and Hogness, D.S. (1983) Cell 34, 807-814.
13. Frielle, T., Collins, S., Daniel, K.W., Caron, M.G., Lefkowitz, R.J., and Kobilka, B.K. (1987) Proc. Natl. Acad. Sci. USA 84, 7920-7924.
14. Kobilka, B.K., Dixon, R.A.F., Frielle, T., Dohlman, H.G., Bolanowski, M.A., Sigal, I.S., Yang-Feng, T.L., Francke, U., Caron, M.G., and Lefkowitz, R.J. (1987) Proc. Natl. Acad. Sci. USA 84, 46-50.
15. Emorine, L.J., Marullo, S., Briend-Sutren, M.M., Patey, G., Tate, K., Delavier-Klutchko, C., and Strosberg, A.D. (1989) Science 245, 1118-1121.

16. McFarland, K.C., Sprengel, R., Phillips, H.S., Kohler, M., Rosemblit, N., Nikolics, K., Segaloff, D.L., and Seeburg, P.H. (1989) Science, 494-499.
17. Takeya, T., and Hanafusa, H. (1983) Cell 32, 881-890.
18. Sukegawa, J., Semba, K., Yamanashi, Y., Nishizawa, M., Miyajima, N., Yamamoto, T., and Toyoshima, K. (1987) Mol. Cell. Biol. 7, 41-47.
19. Katamine, S., Notario, V., Durga Rao, C., Miki, T., Cheah, M.S.C., Tronik, S.R., and Robbins, K. C. (1988) Mol. Cell. Biol. 8, 259-266.
20. Cooper, J.A., Gould, K.L., Cartwright, C.A., and Hunter, T. (1986) Science 231, 1431-1434.
21. Strader, C.D., Sigal, I.S., and Dixon, R.A.F. (1989) FASEB J. 3, 1825-1832.
22. Lohse, M.J., Lefkowitz, R.J., Caron, M.G., and Benovic, J.L. (1989) Proc. Natl. Acad. Sci. USA 86, 3011-3015.
23. Hopp, T.P., and Woods, K.R. (1981) Proc. Natl. Acad. Sci. USA 78, 3824-3828.
24. Cotecchia, S., Schwinn, D.A., Randall, R.R., Lefkowitz, R.J., Caron, M.G., and Kobilka, B.K. (1988) Proc. Natl. Acad. Sci. USA 85, 7159-7163.
25. Regan, J.W., Kobilka, T.S., Yang-Feng, T.L., Caron, M.G., Lefkowitz, R.J., and Kobilka, B.K. (1988) Proc. Natl. Acad. Sci. USA 85, 6301-6305.
26. Peralta, E.G., Ashkenazi, A., Winslow, J.W., Smith, D.H., Ramachandran, J., and Capon, D.J. (1987) EMBO J. 6, 3923-3929.
27. Liao, C.F., Themmen, A.P.N., Joho, R., Barberis, C., Birnbaumer, M., and Birnbaumer, L. (1989) J. Biol. Chem. 264, 7328-7337.
28. Bunzow, J.R., Van Tol, J.H.M., Grandy, D.K., Albert, P., Salon, J., Christie, M.D., Machida, C.A., Neve, K.A., and Civelli, O. (1988) Nature 336, 783-786.
29. Fargin, A., Raymond, J.R., Lohse, M.J., Kobilka, B.K., Caron, M.G., and Lefkowitz, R.J. (1988) Nature 335, 358-360.
30. Julius, D., MacDermott, A.B., Axel, R., and Jessell, T.M. (1988) Science 241, 558-564.
31. Pritchett, D.B., Bach, W.J., Wozny, M., Taleb, O., Dal Tosso, R., Shih, J.C., and Seeburg, P.H. (1988) EMBO J. 7, 4135-4140.
32. Masu, Y., Nakayama, K., Tamaki, H., Harada, Y., Motoy, K., and Nakanishi, S. (1987) Nature 329, 836-838.