Volume 264, number 2, 193-197 FEBS 08423 May 1990

Molecular cloning of a human thyrotropin receptor cDNA fragment

Use of highly degenerate, inosine containing primers derived from aligned amino acid sequences of a homologous family of glycoprotein hormone receptors

Guo Cai Huang[1], Martin J. Page[2], Andrew J. Roberts[1], Afshan N. Malik[1], Hugh Spence[2], Alan M. McGregor[1] and J. Paul Banga[1]

[1] *Department of Medicine, King's College School of Medicine, Denmark Hill, London SE5 8RX, UK* and [2] *Wellcome Research Laboratories, Wellcome Biotech, Langley Park Road, Beckenham, Kent BR3 3BS, UK*

Received 12 March 1990

Autoantibodies to the thyrotropin (TSH) hormone receptor (TSH-R) are present in the sera of patients with thyroid autoimmune disease which are pathogenetic leading to hyperthyroidism of Graves' disease. Considerable interest has been focused on the cloning of the human TSH-R, which has until very recently, proven exceedingly difficult due to the very low receptor level expression on thyroid cells. We have used polymerase chain reaction and highly degenerate, inosine containing oligonucleotides derived from sequence alignments of the transmembrane regions 2 and 7 of a number of G-binding protein receptors including the lutropin/choriogonadotropin (LH/CG) receptors to amplify various cDNAs from human thyroid cDNA. Sequencing analysis of 27 different clones revealed that they fall into eight different groups. The very recent publication of the complete nucleotide sequence of the human TSH-R revealed that one of the groups (GT1) containing seven clones which had been sequenced belong to the human TSH-receptor. The sequence of all 7 GT1 clones was identical and in complete concordance with transmembrane regions 2 and 7 of the published TSH-R sequence. Our results show that by designing oligonucleotides to common transmembrane regions of G-binding proteins where the primers are biased in their sequence to the LH/CG receptors it is possible to amplify the TSH-R receptor sequence.

TSH receptor, human; Polymerase chain reaction; Graves' disease; G-binding protein receptor

1. INTRODUCTION

The receptor for thyroid stimulating hormone (TSH-R), found on the basal surface of the thyroid follicular cells, governs thyroid cell function and growth [1,2]. Physiological activation of the TSH-R by the pituitary hormone thyrotropin (TSH) enhances adenyl cyclase activity leading to increased thyroid hormone synthesis and thyroid cell growth [3]. The human TSH-R (hTSH-R) is the target of attack in autoimmune thyroid disease with aberrant stimulation of the receptor by autoantibodies leading to the hyperthyroidism observed in Graves' disease [4-8].

The TSH-R has not been amenable to biochemical purification or molecular cloning procedures due to the extraordinary low levels of receptor expressed on thyroid cells, although the biochemical properties of the receptor have been studied [8]. TSH belongs to the glycoprotein family of hormones, which consists of two heterodimeric polypeptide subunits and also includes lutropin (LH), choriogonadotropin (CG) and follitropin (FSH). The α subunit is common to all these hormones, of which each has a distinct β subunit component [9]. Due to this homology it is likely that the receptors for this glycoprotein hormone family should also belong to a common family. Since the biological effects of the glycoprotein hormone family are mediated by cAMP, the receptors belong to the family of G-protein coupled receptors [10,11]. Other members of this family whose sequences have been deduced by cDNA cloning include receptors for the α1, α2, β1 and β2 adrenergic and muscarinic cholinergic receptors, serotonin, substance K, angiotensin, dopamine and the opsin protein, rhodopsin [11]. To this group of membrane proteins which traverse the membrane 7 times, a hallmark for guanine nucleotide binding receptors, have been added the complete sequences of the porcine and rat LH/CG receptors [12,13].

Recent technological advances in the polymerase chain reaction (PCR) methodology first described by Saiki et al [14] have resulted in two major developments which allow amplification and cloning of low abundance mRNAs. Firstly, the ability to use degenerate oligonucleotides derived from peptide sequences for amplification of DNA from total cDNA (termed 'mixed oligonucleotide primed amplification of cDNA (MOPAC)') [15] and, secondly, the amplification of rare transcripts of cDNA using homopolymer tailing

Correspondence address: J.P. Banga, Department of Medicine, King's College School of Medicine, Denmark Hill, London SE5 8RX, UK

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50 © 1990 Federation of European Biochemical Societies

(termed 'rapid amplification of cDNA ends (RACE)') has been described [16]. In addition, Libert et al. have independently used amplification techniques incorporating MOPAC technology together with low stringency hybridisation conditions to identify members of the G-protein coupled receptors by designing highly degenerate oligonucleotides to the common transmembrane regions of the guanine nucleotide binding proteins [17]. Additionally, these primers incorporate inosine [18]. Using this methodology on cDNA prepared from human thyroid gland, with highly degenerate consensus oligonucleotide primers to transmembrane regions 3 and 6, four new members of the G-protein coupled family of receptors were described. However, the transcript for the elusive thyrotropin receptor failed to be amplified with these degenerate primers [17]. During the course of the above experiments, the sequences for the LH/CG receptors [12,13] were not available.

Following the publication of the LH/CG receptors, we have attempted to clone the human TSH-R by employing a combination of the MOPAC technology [15] and a modification of the method described. By aligning the amino acid sequences of G-protein coupled receptors (including the LH/CG receptors [12,13]), we have designed degenerate, inosine containing oligonucleotides which are different from those described in [17] for amplification of the TSH-R cDNA fragment from human thyroid mRNA.

During the course of this work, the complete sequences of dog and human TSH-R were simultaneously published by 3 independent groups [19–23]. The publication of the hTSH-R cDNA sequences [22,23] enabled us to compare the sequences of the cDNA fragments which we had amplified directly from human thyroid cDNA, using highly degenerate, inosine containing oligonucleotides. This allowed us to identify our cDNA clones without any functional expression data. We now report that using a modification of the method originally described [17], it is possible to directly amplify the hTSH-R from human thyroid cDNA. This may have applications to identifying other rare, uncharacterised cDNA transcripts which belong to families of proteins in which some family member sequences have been determined.

2. MATERIALS AND METHODS

2.1. *Synthesis of degenerate and inosine containing oligonucleotides*

Oligonucleotides were synthesised on an automated synthesiser (Milligen 7500 DNA synthesizer) using β-cyanoethylphospharamidite chemistry. Inosine additions were carried out using cyanoethyl deoxyinosine phosphoramidite where 0.25 g was dissolved in 3 ml anhydrous acetonitrile and placed in the 1 position of the synthesiser. Reactions were performed as for normal phosphoramidites but because of the instability of the inosine solution this was used within 2 days of preparation. Mixed base additions were carried out using mixing protocols supplied with the synthesiser or, in the case of mixed 3′ termini, by mixing the individual supports when packing the synthesis columns. The mixed oligonucleotides thus prepared were used subsequently without any further purification by resuspending at 10 pmol/μl in sterile water for PCR.

2.2 *Polymerase chain amplification (PCR)*

Total cellular RNA from a Graves' thyroid gland was isolated by a modification of the single step acid guanidium method [24] and poly A^+ RNA prepared [25]. Double stranded cDNA was synthesised from 5 μg poly A^+ RNA as template and oligo d(T) for first strand synthesis with Moloney Murine Leukemia virus reverse transcriptase using a commercial kit (Pharmacia). The resulting double stranded cDNA was precipitated in ethanol and after washing resupended in 20 μl sterile water.

PCR was performed in a thermal cycler (Perkin Elmer Cetus) with 2.5U Taq polymerase (Cetus) using a modification of the method described [17]. Double stranded cDNA (1 μl) in 100 μl total volume of amplification buffer [14] containing a final concentration of 2 mM Mg^{2+} and 40 pmol of each oligonucleotide primer were subjected to 25 cycles of PCR. Each cycle consisted of 40 s denaturation at 92°C, 1.5 min annealing at 55°C and 3 min chain extension at 72°C. Of the above reaction mixture 2 μl were then subjected to an additional 30 cycles of amplification under identical conditions. Following a total of 55 cycles, the overlying oil was removed and a 10 μl sample was electrophoresed in 1.6% agarose gel to visualize the PCR products. The remainder of the amplified DNA was extracted with phenol/chloroform, ethenol precipitated and following resuspension in 100 μl TE8, digested with *Bam*HI adn *Eco*RI. Following digestion, the amplified DNA was electrophoresed in low melting agarose gel, the 700 bp DNA fragment excised, eluted and purified on an Elutip-d column (Schleicher & Schull).

2.3. *Nucleotide sequencing and comparison of sequences*

The purified 700 bp amplified DNA was ligated into *Bam*HI/*Eco*RI digested M13 (mp18) or pBluescript SKII+ and sequenced by the chain termination method using a sequenase kit (US Biochemicals) and (α^{35}S)dATP (Amersham).

Nucleotide sequences were compared using Wordsearch Program on the GenEMBL databank and compared with each other using the Bestfit Program. Both programs were from the GCG package [26].

3. RESULTS AND DISCUSSION

The complete amino acid sequences of the rat and porcine LH/CG receptors have been recently deduced [12,13]. The transmembrane regions showing the highest degree of homology with other G-protein coupled receptors were the putative transmembrane segments 2, 3, 6 and 7 [12,13,27]. Highly degenerate and inosine containing oligonucleotides were designed to these regions of the putative transmembrane segments of the aligned G-protein coupled receptors. Additionally, the design of the oligonucleotide primers was biased towards the transmembrane segments of the LH/CG receptors [12,13]. The final sequence of the degenerate, inosine-containing oligonucleotide primers derived from the alignments to the putative transmembrane segments 2, 3, 6 and 7 are shown in Fig. 1 and were termed TM2, TM3, TM6 and TM7 primers, respectively. A *Bam*HI restriction site in forward primers TM2 and TM3 and an *Eco*RI restriction site in the reverse primers TM6 and TM7 were incorporated in the 5′ region to facilitate subcloning.

Amplification of different combinations of forward and reverse primers from thyroid double stranded cDNA by PCR gave a specific, amplified fragment of

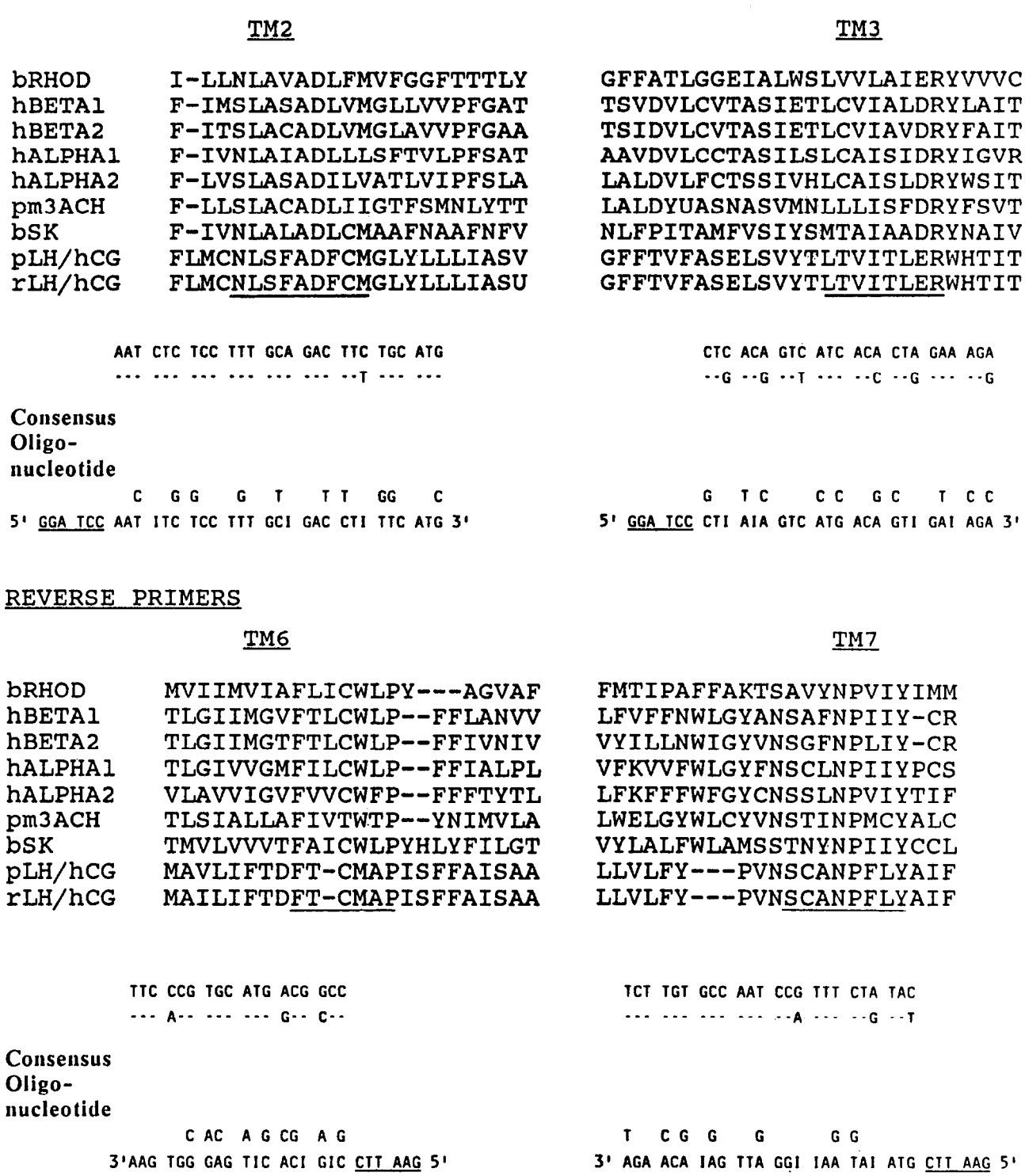

Fig.1. Design of consensus oligonucleotides for amplification of TSH-R cDNA, derived from alignments of transmembrane regions of different G-protein coupled receptors. These are bRHOD, bovine rhodopsin; hBETA1, human β1 adrenergic receptor; hBETA2, human β2 adrenergic receptor; hALPHA1, human α1 adrenergic receptor; hALPHA2, human α2 adrenergic receptor; pm3ACH, porcine muscarinic cholinergic receptor; bSK, bovine substance K receptor; pLH/hCG, porcine lutropin/human choriogonadotropin receptor (derived from [13]); rLH/hCG, rat lutropin/human choriogonadotropin receptor (derived from [12]). The aligned regions represent transmembrane regions 2, 3, 6 and 7 and are termed TM2, TM3, TM6 and TM7, respectively. The amino acids underlined in the alignments represent some of the homologous residues selected for the consensus oligonucleotides. The nucleotide sequences below these amino acids represent the sequence of the cDNA at that region of porcine LH/hCG (upper) and rat LH/hCG (below) receptors. The dashes (---) represent aligned identical nucleotides. The consensus oligonucleotide shows the design of the mixed probe (degenerate) and inosine containing oligonucleotide derived from the alignment. The underlined nucleotides comprise the restriction enzyme sites; a *Bam*HI site in the forward primers TM2 and TM3 and an *Eco*RI site in the reverse primers TM6 and TM7.

the correct size of approximately 700 bp using primers TM2 and TM7 only (Fig. 2). The use of other primer combinations TM2/TM6, TM3/TM6 and TM3/TM7 did not yield any specific amplified products (not shown). Complex amplification protocols from cDNA using degenerate, inosine containing primers are exceedingly susceptible to minute changes in Mg^{2+} concentration, template and primer concentrations and other variables such as the enzyme concentration. Our attempts to amplify any product under a variety of conditions using the latter combinations of primers have been unsuccessful. Additionally, the use of all combinations of the primers described above for RACE PCR[16] amplification of TSH-R have also been unsuccessful.

The amplified DNA fragments using primers TM2 and TM7 were sequenced to identify the G-protein receptor family. Twenty seven different clones containing the amplified cDNAs were sequenced. This revealed that the amplified cDNAs fall into 8 different groups (Table I). Groups GT1 (7 clones) and GT2 (3 clones) were in complete concordance within the transmembrane regions 2 and 7 to the hTSH-R and β2-adrenergic receptor family, respectively [22,23,28]. Group GT4 (4 clones) showed a limited homology with human early growth response factor 2 (EGR-2) [29] and the Krox-20 protein containing zinc fingers [30]. Groups GT3 and GT5–GT8 did not show any homology to any of the gene sequences in the database.

The publication of the complete nucleotide sequences of the dog and the human TSH-R allowed us to compare the sequences of our amplified products to the published TSH-R sequences. Parmentier et al. used degenerate oligonucleotides spanning transmembrane regions 2 and 7 to amplify material from human genomic DNA [19]. Using an amplified cDNA fragment (which they now believe to be FSH-R [19]) as a probe, to screen a dog cDNA library at low stringency, they unambiguously identified two related dog TSH-R cDNAs [19,20]. Screening of a human thyroid cDNA library with dog TSH-R cDNA as a probe led to the isolation of the human hTSH-R [21]. Using a different strategy, based either upon the use of a full length cDNA probe to the LH/CG receptor [23] or oligonucleotides derived from that sequence [22], allowed two independent groups to identify a cDNA clone representing the hTSH-R from a human thyroid cDNA library [22,23]. Interestingly and importantly, all three cDNA clones to the hTSH-R are identical in amino acid sequence [21–23].

Our results show that using degenerate oligonucleotide primers designed according to the strategies already described [17], a rare cDNA transcript such as the TSH-R can be specifically amplified and cloned. The availability of the full nucleotide sequence of the hTSH-R [22,23] allowed us to unambiguously confirm 7 of the 27 amplified cDNA clones to belong to the hTSH-R gene, without the need to identify a full length clone and functional expression data. Interestingly, a similar cloning strategy using degenerate oligonucleotide primers to transmembrane regions 3 and 6 described in [17], has allowed the cloning of an alternatively spliced member of another G-protein coupled receptor, the dopamine D_2 receptor of rat brain [31]. Additionally amplification of unique members of the protein tyrosine kinase family has also been described using degenerate oligonucleotides derived from highly conserved sequence motifs of members of this family [32]. The method described herein and

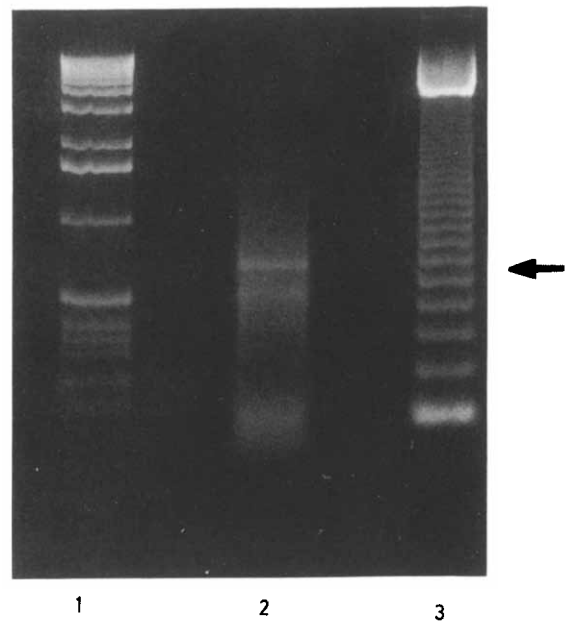

Fig.2. Ethidium bromide stained agarose gel to show in lane 2 the 700 bp amplified cDNAs (arrowed) by a PCR from a human thyroid double stranded cDNA. The consensus oligonucleotide primers TM2 and TM7 were used. Lanes 1 and 3 are the 1 kb and the 123 bp ladders (BRL).

Table I

Derivation of the twenty seven cDNA clones amplified with TM2 and TM7 primers into eight distinct groups and their homologies where they exist with the published sequences.

Group	Number of clones	Homology
GT1	7	hTSH-R [22,23]
GT2	3	β2 adrenergic receptor [27]
GT3	6	unknown
GT4	4	limited similarity with early growth response gene-2 (EGR-2) [29] and Krox-20 protein containing zinc finger [30]
GT5	3	unknown
GT6	1	unknown
GT7	2	unknown
GT8	1	unknown

The hTSH-R amplified clone of group GT1 was fully sequenced whilst all the remaining clones were partially sequenced (minimum 250 bp from 670 bp amplified DNA).

```
          1463                                                    1513
hTSH-R    5' AACCTGGCCTTTGCGGATTTCTGCATGGGGATGTACCTGCTCCTCATCGCC.......
             1                                                      50
GT1.1     5' AATCTCGCCTTTGTGCATTTGTGCATGGGGATGTACCTGCTCCTCATCGCC.......
to GT1.7
```

```
...... 2083                                                    2133
          TTGCTGGTACTCTTCTATCCACTTAACTCCTGTGCCAATCCATTCCTCTAT 3' hTSH-
.........620                                                    670
          TTGCTGGTACTCTTCTATCCACTTAACTCTGGCCCCAACCCCCTTATCTAC 3'GT1.1
                                                               to GT1.7
```

Fig.3. Nucleotide sequence of the GT1 clones amplified from thyroid cDNA with TM2 and TM7 primers. The sequence of all 7 GT1 clones termed GT1.1 to GT1.7 were identical. The nucleotide sequence of the 670 bp GT1 clones is shown from residue 1 to 50 and from 620 to 670. Directly above is the nucleotide sequence of the published human TSH-R [22,23] from residue 1463 to 1513 and 2083 to 2133 (encompassing transmembrane regions 2 and 7). The underlined nucleotides in the GT1 clones represent the primer sequences used for amplification. The sequences between the primer residues in GT1 clones were identical in sequence to the published TSH-R sequence.

those of [17,32] are applicable to cloning other rare cDNA transcripts belonging to families of other related proteins for which there may be some sequence information.

The nucleotide sequences of the groups GT3 and GT5 to GT8 are being deposited at the EMBL Database Library.

Acknowledgements: G.C.H. was supported by a Royal Society and British Academy K.C. Wong Fellowship for China. We wish to thank Dr R.W.S. Tomlinson and P. Cunningham for assistance with the computer analysis and Mrs J. De Groote for preparation of the manuscript. This work was supported by grants from The Wellcome Trust.

REFERENCES

[1] Dumont, J.E. (1971) Vitam. Horm. 29, 287–406.
[2] Roger, P.P. and Dumont, J.E. (1984) Mol. Cell. Endocrinol. 36, 79–93.
[3] Roger, P.P., Hotimsky, A., Moreau, C. and Dumont, J.E. (1982) Mol. Cell. Endocrinol. 26, 165–170.
[4] Adams, D.D. (1956) J. Clin. Endocrinol. Metab. 18, 699–712.
[5] Dorrington, K.J. and Munro, D.S. (1966) Clin. Pharmacol. Ther. 7, 788–794.
[6] McKenzie, J.M. (1968) Physiol. Rev. 43, 252–271.
[7] Smith, B.R. and Hall, R. (1974) Lancet 2, 427–431.
[8] Rees-Smith, B., McLachlan, S.M. and Furmaniak, J. (1988) Endocr. Rev. 9, 106–121.
[9] Ryan, R.J., Charlesworth, M.C., McCormick, D.J., Milius, R.P. and Keutmann, H.T. (1988) FASEB J. 2, 2661–2669.
[10] Schulz, S., Chinkers, M. and Garbers, D.L. (1989) FASEB J. 3, 2026–2035.
[11] Gilman, A.G. (1987) Annu. Rev. Biochem. 56, 615–624.
[12] McFarland, K.C., Sprengel, R., Phillips, H.S., Kohler, M., Rosenblit, N., Nicolics, K., Segaloff, D.L. and Seeburg, P.H. (1989) Science 245, 494–499.
[13] Loosfelt, H., Misrachi, M., Atger, M., Salesse, R., Thi, M.T., Jolivet, A., Guiochan-Mantel, A., Sar, S., Jallal, B., Garnier, J. and Milgrom, E. (1989) Science 245, 525–528.
[14] Saiki, R.K., Bugawan, T.L., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1986) Nature 324, 163–166.
[15] Lee, C.C., Wu, X.W., Gibbs, R.A., Cook , R.G., Muzny, D.M. and Caskey, C.T. (1988) Science 239, 1288–1291.
[16] Frohman, M.A., Dush, M.K. and Martin, G.R. (1988) Proc. Natl. Acad. Sci. USA 85, 8998–9002.
[17] Libert, F., Parmentier, M., Lefort, A., Dinsart, C., Van Sande, J., Maenhavt, C., Simons, M.J. and Vassart, G. (1989) Science 244, 569–572.
[18] Knoth, K., Roberts, S., Poteet, C. and Tomkin, M. (1988) Nucl. Acids Res. 16, 10932.
[19] Parmentier, M., Libert, F., Maenhaut, C., Lefort, A., Gerard, C., Perret, J., Van Sande, J., Dumont, J.E. and Vassart, G. (1989) Science 246, 1620–1622.
[20] Parmentier, M., Libert, F., Maenhaut, C., Lefort, A., Gerard, C., Perret, J., Van Sande, J., Dumont, J.E. and Vassart, G. (1989) Nucl. Acids Res. 17, 10493.
[21] Libert, F., Lefort, A., Gerard, C., Parmentier, M., Perret, J., Ludgate, M., Dumont, J.E. and Vassart, G. (1989) Biochem. Biophys. Res. Commun. 165, 1250–1255.
[22] Nagayama, Y., Kaufmann, K.D., Seto, P. and Rapoport, B. (1989) Biochem. Biophys. Res. Commun. 165, 1184–1190.
[23] Misrahi, M., Loosfelt, H., Atger, M., Sar, S., Guiochon-Mantel, A. and Milgrom, E. (1990) Biochem. Biophys. Res. Commun. 166, 394–403.
[24] Kung, A.W.C., Collison, K., Banga, J.P. and McGregor, A.M. (1988) FEBS Lett. 232, 12–16.
[25] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: a Laboratory Manual, Cold Spring Harbour Laboratory, Cold Spring Harbor, NY.
[26] Devereux, J., Haeberli, J. and Smithies, O. (1984) Nucl. Acids Res. 12, 387–395.
[27] 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.
[28] Kobilka, B.K., Dixon, R.A.F., Frielle, T., Dohlman, H.G. et al. (1987) Proc. Natl. Acad. Sci. USA 84, 46–50.
[29] Joseph, L.J. Beau, M.M., Jamieson, J., Acharya, S. et al. (1988) Proc. Natl. Acad. Sci. USA 85, 7164–7168.
[30] Chavrier, P., Zerial, M., Lemaire, P., Almendral, J., Bravo, R. and Charnay, P. (1988) EMBO J. 7, 29–35.
[31] Eidne, K.A., Taylor, P.L., Zebavnik, J., Saunders, P.T.K. and Inglis, J.D. (1989) Nature 342, 865.
[32] Wilks, A.F., Kurban, R.R., Hovens, C.M. and Ralph, S.J. (1989) Gene 85, 67–74.