Isolation and characterization of a cDNA clone for porcine thyroid peroxidase

Ronald P. Magnusson, Jane Gestautas, Pui Seto, Alvin Taurog and Basil Rapoport

Departments of Medicine, Metabolism (111F), VA Medical Center, San Francisco, CA 94121 and the University of California, San Francisco; and the Department of Pharmacology, University of Texas Health Science Center, Dallas, TX, USA

Received 6 October 1986

We undertook the molecular cloning of porcine thyroid peroxidase (TPO). Four oligonucleotide probes were synthesized on the basis of amino acid sequences of 3 tryptic peptides from highly purified porcine TPO. These probes were used to screen a pig thyroid cDNA library. Seven of 16 selected clones (0.45–1.15 kb in size) reacted with all 4 probes. Nucleotide sequencing of the 1.15 kb at the 3′-end of the structural gene revealed the complementary sequence to all 4 probes as well as the nucleotides coding for the entire length of the 3 tryptic peptides. There is an open reading frame of 332 amino acid residues. On Nothern blot analysis this gene codes for an mRNA species of 2.85 kb, corresponding to the anticipated size of the mRNA for the intact TPO molecule. We have therefore cloned and characterized a cDNA clone coding for approx. 36% of porcine thyroid peroxidase.

Thyroid Peroxidase cDNA clone

1. INTRODUCTION

Thyroid peroxidase (TPO) is a hemoprotein whose purification has been difficult because it is membrane-bound and therefore requires solubilization with detergent and limited trypsin digestion [1–5]. For these reasons there has been considerable disagreement in the literature as to some of its basic properties. For example, estimates of its molecular mass have varied between 45 and 104 kDa (discussed in [1,6]). Recent purification of TPO using monoclonal antibodies without tryptic digestion suggests that in its native state the enzyme is present as a single polypeptide of about 100 kDa and may exist as a 400 kDa tetramer [7].

The molecular cloning of TPO would represent a major advance in efforts to characterize definitively this important thyroid enzyme. Here, we report the cloning and characterization of a 1.15 kb cDNA clone for the 3′-end of porcine TPO.

2. MATERIALS AND METHODS

2.1. *Construction of pig thyroid cell cDNA library*

The oligo(dT)-primed pig thyroid cell cDNA library was constructed in λ gt11 essentially as described by Ebina et al. [8], with minor modifications, using mRNA prepared [9,10] from primary cultures of pig thyroid cells stimulated with 25 mU/ml TSH for 2 days. Double-stranded cDNA was ligated into the *Eco*RI site of λ gt11 [11] (Vector Cloning System, San Diego, CA) with a yield of 1.26×10^6 recombinant clones, followed by amplification in Y1088 cells [11]. The average size insert ($n = 10$) was 0.93 kb.

2.2. *Oligonucleotide probe synthesis and screening*

Preparation of the purified porcine TPO used for amino acid analysis was described previously [4]. Estimated purity was 71–84% based on a comparison of the A_{410}/A_{280} ratio to a preparation shown by polyacrylamide gel electrophoresis to be

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

80–95% pure. 100 μg was provided to Dr Craig Miles of the National Jewish Center for Immunology and Respiratory Medicine, Denver, CO, who performed the amino acid analysis, tryptic digestion, HPLC purification of these fragments and amino acid sequencing of selected fragments [12]. The oligonucleotide probes were synthesized according to our specifications by the Biomolecular Resource Center, University of California, San Francisco.

The oligonucleotide probes were end-labeled with [γ-^{32}P]ATP (New England Nuclear, Boston, MA) to a specific activity of about 1–2 μCi/pmol [13]. Screening was performed according to standard techniques [13]. Hybridizations were conducted for 2–16 h at 42°C in 6 × SSC buffer (150 mM NaCl, 15 mM Na citrate, pH 7.0), 10 mM EDTA, 5 × Denhardt's solution, 0.5% Na lauryl sarcosine, 0.1 mM ATP and 1.0 mg/ml tRNA. Final rinses (20 min each) were at 42°C (twice) and at 48°C (once) in 6 × SSC, 0.1% Na lauryl sarcosine. Autoradiographs were performed at −70°C using Kodak XAR-5 film.

2.3. *Subcloning and DNA sequencing*

cDNA inserts were released by *Eco*RI digestion (10 U/μg DNA; Boehringer Mannheim, Indianapolis, IN) and ligated into *Eco*RI-cut pUC18 or M13. Dideoxynucleotide sequencing was performed using 4–6% denaturing polyacrylamide urea gels [14].

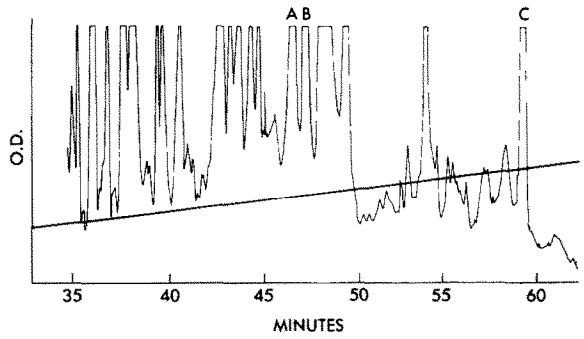

Fig.1. Tryptic peptides of purified pig thyroid TPO. 100 μg purified TPO was subjected to tryptic digestion followed by HPLC [16]. Three hydrophobic peptides (A eluting at 47.02 min, B at 47.70 min, and C at 59.67 min) were chosen for amino acid sequence analysis.

2.4. *Northern blot analysis*

10 μg poly(A$^+$)-mRNA from pig thyroid tissue was electrophoresed in 1% agarose using formaldehyde [13], blotted onto nitrocellulose paper and probed with a [^{32}P]dCTP nick-translated porcine TPO cDNA.

3. RESULTS

3.1. *Cloning of TPO recombinants using oligonucleotide probes*

We first screened the λ gt11 expression library using rabbit antiserum to porcine TPO prepared

Probe[A]	Peptide[B]	Amino Acid[C]	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	B		Phe	Trp	Trp	Glu	Asn	Pro	Gly	Val	Phe	Thr	Glu	Ala	Gln	Arg			
			AA^G_A	ACC	ACC	CT^T_C	TT^G_A	GG(N)	CC										
2	A		Val	Gly	Gln	Trp	Pro	Gln	Glu	Phe	Glu	Pro	Cys	Ala	Ser	Ile	Gln	Gly	Met
					GT^T_C	ACC	GG(N)	GT^T_C	CT^T_C	AA^G_A	CT^T_C	GG							
3			GTG	GGC	CAG	TGG	CCC	CAG	GAG	TTC	GAG	CCC	TGC	GCC	TCC	ATC	---	GGC	ATG
4	C		Ile	Leu	Gly	Leu	Tyr	Gln	His	Pro	(Bla)	Asn	Ile	(Bla)	Val	Trp	Leu		
			TAG	GAC	CCG	GAC	AT^G_A	GT^T_C	GT^G_A	GG									

Fig.2. Oligonucleotide probes synthesized on the basis of TPO tryptic peptides A–C (see fig.1). Probes 1, 2 and 4 are anticodon sequences; probe 3 the codon sequence. Two nucleotides at one position indicate that both were synthesized. BLA (Blank) indicates that this residue could not be identified, probably because of derivatization.

previously by one of us (A.T.). However this approach was unsuccessful because all clones also reacted with preimmune serum, though not with normal rabbit serum. We therefore turned to synthesizing oligonucleotide probes. After tryptic digestion of highly purified porcine TPO three hydrophobic peptides from the HPLC profile (A–C) were chosen on the basis of their strength of signal and good separation from adjacent peaks (fig.1). Based on these peptides, 4 oligonucleotide probes were synthesized (fig.2) according to the following strategies. Probe 1 (20-mer) and probe 2 (23-mer) consisted of all possible permutations (32- and 128-fold degeneracies, respectively). For probe 3, we synthesized a single species of greater length (48-mer) using the nucleotide of greatest frequency in degenerate codons [15]. Probe 4 (23-mer) was synthesized using a combined approach. Thus for 4 amino acid residues, 11 nucleotides containing all possible permutations were selected (degeneracy of 8). For the remaining 4 amino acids the nucleotides of greatest frequency were chosen.

Screening 1.1×10^5 recombinant plaques in the library with probe 4 yielded 20 clones. *Eco*RI digests of DNA from 12 of these clones were subjected to Southern blot analysis (fig.3), confirming the specificity of the interaction of the oligonucleotide probe with the cDNA inserts. Inserts in clones D, H and P apparently had two fragments, representing either an *Eco*RI site, or the cloning of more than one cDNA fragment into the same vector. As expected, only one fragment reacted with probe in these clones. To test whether probes 1–3 would hybridize to probe 4-selected clones, and also to obtain information on the possible interrelationships between these clones, 16 clones were tested in a confluent plaque spot assay

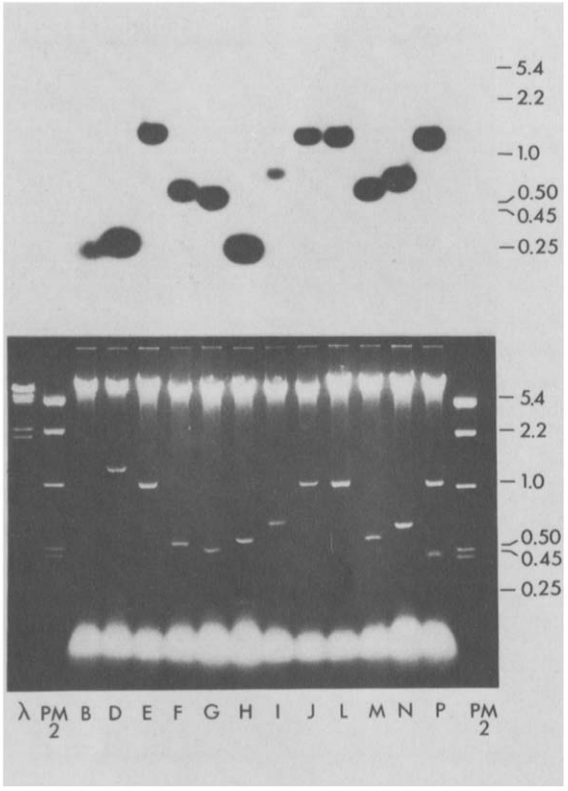

Fig.3. Agarose gel electrophoresis and Southern blot analysis of *Eco*RI digests of probe 4 positive clones. Phage DNA from 12 of the probe 4 positive clones (B–P) was digested with *Eco*RI and electrophoresed in 1.2% agarose, along with *Hin*dIII digests of wild-type bacteriophage λ DNA and bacteriophage PM2 DNA as molecular mass markers. After staining with ethidium bromide for visualization (lower panel), the DNA was blotted onto nitrocellulose paper, hybridized with probe 4 as described in section 2, and autoradiographed (upper panel).

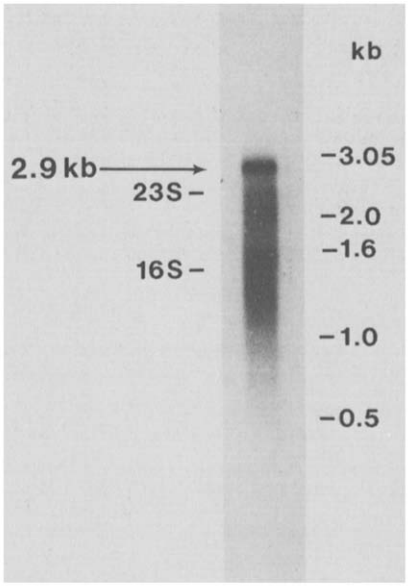

Fig.4. Northern blot analysis of pig thyroid cell mRNA (10 μg) using the cDNA insert (0.6 kb) of clone M as a probe. Molecular size markers (1 kb ladder, BRL, Gaithersburg, MD) are shown on the right, and 16 S and 23 S ribosomal markers are shown on the left.

(A)

```
                                                                                                                  AA
1                                    10                                          20                                          30
Phe Cys Gly Leu Ser Arg Leu Glu Thr Trp Ala Asp Leu Ser Ala Ala Thr Ala Asn Gly Arg Val Ala Asp Arg Ile Leu Gly Leu Tyr
TTC TGC GGC CTG TCC AGA CTC GAG ACC TGG GCC GAC CTG AGT GCT GCC ACT GCC AAC GGG CGT GTG GCC GAC AGG ATC CTG GGC CTG TAC

                                     40                                          50                                          60
Gln His Pro Asp Asn Ile Asp Val Trp Leu Gly Gly Leu Ala Glu Ser Phe Leu Pro Gly Ala Arg Thr Gly Pro Leu Phe Ala Cys Ile
CAG CAT CCG GAT AAC ATT GAC GTC TGG CTG GGC GGC TTG GCC GAG AGC TTC CTC CCT GGG GCG CGG ACC GGC CCG CTG TTC GCC TGC ATC

                                     70                                          80                                          90
Ile Gly Lys Gln Met Arg Ala Leu Arg Asp Gly Asp Arg Phe Trp Trp Glu Asn Pro Gly Val Phe Thr Glu Ala Gln Arg Arg Glu Leu
ATC GGA AAG CAG ATG AGG GCC CTG AGG GAC GGC GAC CGG TTC TGG TGG GAG AAC CCG GGG GTG TTC ACA GAA GCC CAG AGG CGC GAG CTG

                                     100                                         110                                         120
Ser Arg His Ser Met Ser Arg Val Ile Cys Asp Asn Ser Gly Leu Ser His Val Pro Leu Asp Ala Phe Arg Val Gly Gln Trp Pro Gln
AGC CGG CAC TCT ATG TCC CGC GTC ATC TGC GAC AAC AGC GGC CTG TCC CAC GTG CCC CTT GAT GCC TTC CGG GTG GGC CAG TGG CCT CAG

                                     130                                         140                                         150
Glu Phe Glu Pro Cys Ala Ser Ile Gln Gly Met Asp Leu Gly Ala Trp Arg Glu Ala Pro Pro Ser Gly Asp Ala Cys Gly Phe Pro Asp
GAG TTC GAG CCG TGT GCC AGC ATC CAG GGC ATG GAC CTG GGC GCG TGG AGG GAG GCC CCT CCG TCA GGG GAC GCG TGT GGC TTC CCG GAC

                                     160                                         170                                         180
Pro Val Glu Asp Gly Gly Phe Leu Leu Cys Glu Glu Arg Gly Gln Arg Val Leu Val Phe Ser Cys Arg His Gly Phe Arg Leu Arg Gly
CCA GTG GAA GAC GGG GGC TTC CTG CTC TGT GAG GAG CGT GGC CAG CGC GTG CTG GTG TTT TCC TGT CGT CAC GGG TTC CGG CTC CGA GGA

                                     190                                         200                                         210
Pro Ala Gln Ile Thr Cys Thr Pro Arg Gly Trp Asp Ser Pro Pro Pro Leu Cys Lys Asp Ile Asn Glu Cys Glu Asp Glu Thr Asp Pro
CCA GCG CAG ATC ACC TGC ACC CCC CGA GGG TGG GAC TCC CCG CCC CCC CTC TGT AAA GAC ATC AAC GAG TGT GAA GAC GAG ACA GAC CCT

                                     220                                         230                                         240
Pro Cys His Ala Ser Ala Arg Cys Lys Asn Thr Lys Gly Gly Val Leu Cys Glu Cys Ser Asp Pro Leu Val Leu Gly Glu Asp Gly Arg
CCC TGC CAC GCG TCT GCC CGG TGC AAG AAC ACC AAG GGT GGC GTC CTG TGC GAG TGC TCG GAC CCT CTC GTG CTC GGG GAG GAC GGC AGG

                                     250                                         260                                         270
Thr Cys Val Asp Ala Gly Arg Leu Pro Arg Ala Ser Val Val Ser Ile Ala Leu Gly Ala Val Leu Val Cys Gly Leu Ala Gly Leu Ala
ACC TGC GTG GAT GCC GGG AGG CTC CCG CGG GCG TCT GTG GTC TCC ATC GCG CTG GGC GCC GTG CTC GTC TGC GGC CTC GCA GGC CTC GCC

                                     280                                         290                                         300
Trp Thr Val Val Cys Arg Trp Thr His Ala Asp Ala Arg Pro Leu Leu Pro Val Gly Glu Gly Glu Gly Asp Gly Lys Ser Pro Ser Leu
TGG ACG GTG GTT TGC AGG TGG ACA CAC GCG GAT GCC AGG CCC TTG CTG CCC GTC GGG GAG GGA GAA GGA GAC GGG AAA AGC CCC TCC CTG

                                     310                                         320                                         330
Pro Leu Pro Gly Cys Gly Asn Arg Arg Asp Val Gly Ala Ala Pro Ala Leu Glu Val Glu Gln Asp Leu Ser Cys Gly Ser Arg Gly Leu
CCG CTG CCC GGA TGC GGG AAC CGC CGG GAT GTG GGC GCT GCT CCC GCT CTG GAG GTG GAG CAG GAC CTG AGC TGT GGA TCC CGA GGC CTC

Cys Glu
TGC GAG TAGGAACCAGCCTGTGTGCCACACGTGCTGCCCGCAGGGTCAGAGTCAGATGCCACCCGCGTGCGTTCCTTTCACAGGCCAGGGAGCAGTGGGAGGGACGATCAGGGCATC

CGGACAAGCTGGTGGAATAAACCCTGCCTCCTTGCA(n)
```

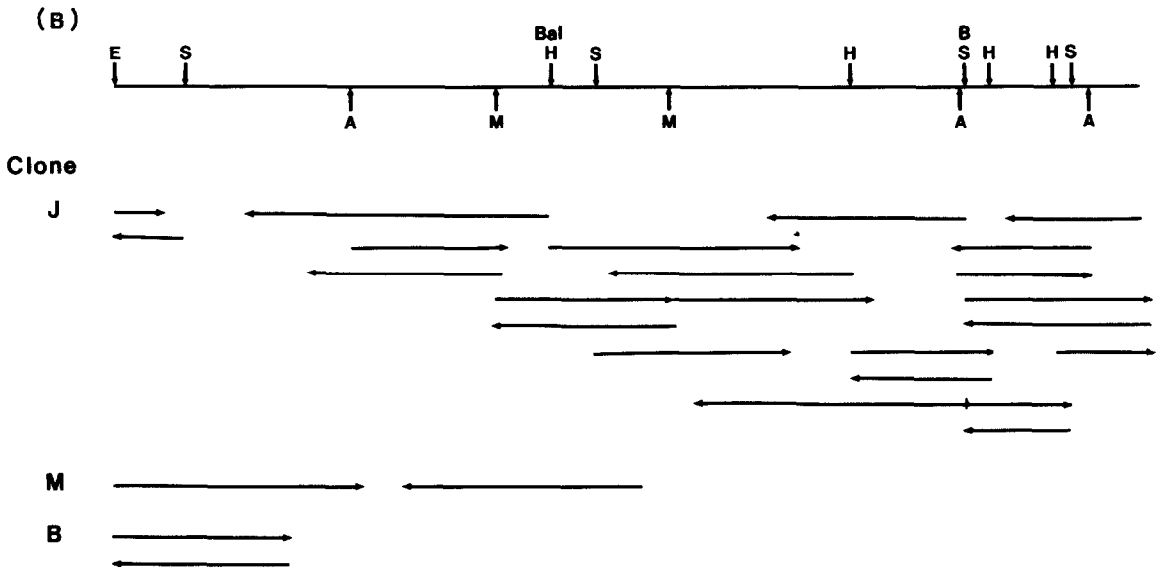

Fig.5. (A) Nucleotide sequence of the 3′-end of porcine TPO cDNA. Peptides A–C are underlined with solid lines. (B) Sequencing strategy employed. Three different clones (J,B,M) were used to obtain the sequence in both directions of this region of the structural gene. E, *Eco*RI; S, *Sau*3A; A, *Alu*I; M, *Mlu*I; Bal, *Bal*I; H, *Hae*III; B, *Bam*HI.

using all 4 probes (not shown). Seven clones reacted with all 4 probes. None of the 4 probes hybridized to any of 5 control clones.

3.2. *Confirmation of TPO cloning by nucleotide sequencing*

The cDNA in one 1.15 kb clone (J) that reacted with all 4 probes was sequenced (fig.5A) according to the strategy shown in fig.5B. Two smaller clones (B, M) which contained identical 5′-ends to clone J but were of shorter length were used to complete the sequencing in both directions of the 1.15 kb at the 3′-end of the structural gene. There is an open reading frame of 332 amino acids followed by an untranslated region of 146 bp prior to the poly(A) tail. The AATAAA polyadenylation signal begins 20 bp upstream of the poly(A) tail. The complementary sequence to all 4 probes was confirmed, and the perfect homology with probe 4 explained the success with this probe. Most important, however, was that the nucleotide sequence flanking the site of hybridization of all of the probes conformed exactly to the anticipated sequence based on the tryptic peptides. The two indeterminate amino acids in peptide C were identified as aspartic acid, which might be anticipated because of their propensity to derivatization. Interestingly, the nucleotide sequence coding for the amino acids in peptide C revealed an unexpected Gln at residue 15 that was not used in the synthesis of probe 3. On reviewing the amino acid sequencing data, the missing Gln cycle was indeed present but had inadvertently not been recorded. Northern blot analysis of porcine thyroid mRNA using the cDNA insert (0.6 kb) of clone M as a probe revealed a single mRNA species of 2.85 kb (fig.4).

4. DISCUSSION

The nucleotide sequence of the 1.15 kb cDNA clone that reacts with all 4 probes indicates unequivocally that we have cloned part of the protein from which the amino acid sequence was derived, namely porcine TPO. The TPO subjected to amino acid sequencing was of high purity (approx. 80%), with no discrete bands other than the enzyme being observed on polyacrylamide gel electrophoresis [4], and the tryptic fragments selected following HPLC separation were discrete.

Further support that the cDNA clone that we have isolated is indeed TPO is provided by the Northern blot analysis that revealed a single mRNA species of 2.85 kb. After subtraction of the nucleotides at the untranslated 3′-end, it can be

deduced that this mRNA species codes for a protein of approx. 107 kDa in size. This size corresponds to the approx. 100 kDa size of intact porcine TPO [7]. An important conclusion from our data, therefore, is that TPO is synthesized as a single peptide. The 'subunits' previously observed [4] can probably be attributed to cleavage by trypsin during the purification of the enzyme.

The nucleotide sequence of 1.15 kb of the 3'-end of porcine TPO reveals an open reading frame of 332 amino acid residues. This would account for a polypeptide of about 36 kDa, approx. 36% of the native enzyme [7]. None of the clones that we isolated from our library contain the entire structural gene. This is not too surprising considering that our library was constructed using an oligo(dT) primer which would bias our recombinants to the 3'-end of the gene. In addition, the average length of our cDNA inserts was only 0.95 kb.

The isolation of a cDNA clone containing part of the porcine TPO structural gene will now permit the cloning and sequencing of the entire TPO gene in the pig, as well as in other species. This knowledge will in turn lead to a better understanding of: (i) the structure and mechanism of action of TPO, and the identification of possible alterations in the TPO genome in various thyroid disorders; (ii) the mechanism by which TSH regulates TPO expression in thyroid cells [16,17]; and (iii) the relationship between TPO and the microsomal antigen in Hashimoto's thyroiditis [18,19].

ACKNOWLEDGEMENTS

This work was supported by NIH grants AM 36182 and 19289 as well as by the Veterans Administration. We greatly appreciated helpful discussions with Drs Walter L. Miller, Bon-Chu Chung, Karla J. Matteson, Charles Craik and Craig Miles. The expert secretarial assistance of Jan Alfstad is gratefully acknowledged.

REFERENCES

[1] DeGroot, L.J. and Niepomniszcze, H. (1977) Metabolism 26, 665–718.
[2] DeGroot, L.J. and Davis, A.M. (1962) Endocrinology 70, 505–510.
[3] Alexander, N.M. (1977) Endocrinology 100, 1610–1620.
[4] Rawitch, A.B., Taurog, A., Chernoff, S.B. et al. (1979) Arch. Biochim. Biophys. 194, 244–257.
[5] Ohtaki, S., Nakagawa, H., Nakamura, M. et al. (1982) J. Biol. Chem. 257, 761–766.
[6] Ohtaki, S., Nakagawa, H., Nakamura, S. et al. (1985) J. Biol. Chem. 260, 441–448.
[7] Nakagawa, H., Kotani, T., Ohtaki, S. et al. (1985) Biochem. Biophys. Res. Commun. 127, 8–14.
[8] Ebina, Y., Ellis, L., Jarnagin, K. et al. (1985) Cell 40, 747–758.
[9] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. et al. (1979) Biochemistry 18, 5294–5299.
[10] Cathala, G., Savouret, J., Mendez, B. et al. (1983) DNA 2, 329–335.
[11] Huynh, T.V., Young, R.A. and Davis, R.A. (1984) in: DNA Cloning Techniques: A Practical Approach (Glover, D. ed.) pp.49–78, IRL, Oxford.
[12] Hannum, C.H., Kappler, J.W., Towbridge, I.S. et al. (1984) Nature 312, 65–67.
[13] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[14] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[15] Grantham, R., Gautier, C., Gouy, M., Jacobzone, M. and Mercier, R. (1981) Nucleic Acids Res. 9, r43–r74.
[16] Magnusson, R.P. and Rapoport, B. (1985) Endocrinology 116, 1493–1500.
[17] Nagasaka, A. and Hidaka, H. (1980) Biochem. Biophys. Res. Commun. 96, 1143.
[18] Portmann, M.D., Hamada, N., Heinrich, G. et al. (1985) J. Clin. Endocrinol. Metab. 61, 1001–1003.
[19] Czarnocka, B., Ruf, J., Ferrand, M., Carayon, P. and Lissitzky, S. (1985) FEBS Lett. 190, 147–152.