Volume 266, number 1,2, 128–132 FEBS 08519 June 1990

The primary structure of the larger subunit of soluble guanylyl cyclase from bovine lung

Homology between the two subunits of the enzyme

Doris Koesling[1], Christian Harteneck[1], Peter Humbert[1], Armin Bosserhoff[2], Rainer Frank[2], Günter Schultz[1] and Eycke Böhme[1]

[1]*Institut für Pharmakologie, Freie Universität Berlin, Thielallee 67–73, D-1000 Berlin 33, FRG* and [2]*European Molecular Biology Laboratory, Postfach 102209, D-6900 Heidelberg, FRG*

Received 10 April 1990; revised version received 27 April 1990

The primary structure of the larger subunit of the soluble guanylyl cyclase from bovine lung, which catalyzes the formation of cyclic GMP from GTP, has been determined. Two clones, isolated from two bovine libraries yielded a total of 3261 bp with a coding region of 2073 bp. The open reading frame encodes a protein of 691 amino acids and a molecular mass of 77500. The deduced amino acid sequence reveals regions which are, to a large extent, homologous to the sequence of the smaller subunit of the enzyme as well as to the sequences of other gyanylyl and adenylyl cyclases.

Soluble guanylyl cyclase; cDNA sequence; Amino acid sequence

1. INTRODUCTION

In most tissues, at least two forms of guanylyl cyclases (GTP pyrophosphate-lyase (cyclizing); EC 4.6.1.2) exist; the soluble and the plasma membrane-bound enzyme forms differ in structure, regulation and other properties [1,2]. Their enzymatic product, cGMP, acts as an intracellular messenger, activating cGMP-dependent protein kinases [3] and regulating cGMP-sensitive ion channels [4]. Although the role of cGMP is generally not well understood, cGMP is an established mediator at least in vascular smooth muscle relaxation and retinal phototransduction [1]. The plasma membrane-bound forms of guanylyl cyclase known so far exist as single polypeptide chains and are regulated by peptide hormones binding to an extracellular receptor domain of the protein [5]. On the other hand, the soluble enzyme form is a heterodimer, which is regulated by nitric oxide (NO) and NO-containing compounds [1,2]. This enzyme, which consists of two subunits, has been purified by several laboratories to apparent homogenity; the reported M_r of the larger subunit varied between 73 and 82 kDa on SDS-polyacrylamide gels [13,14]. The primary structures of several plasma membrane-bound guanylyl cyclases and of the smaller (70 kDa) subunit of the soluble enzyme have been determined [5–11]. Comparisons of these sequences revealed a highly homologous region (about 250 amino acids) in the C-terminal part of the sequences [5]. This homologous region is also found in the two hydrophilic domains of the brain adenylyl cyclase [12].

Here, we report the cDNA cloning and sequencing of the larger subunit of soluble guanylyl cyclase, which shows overall homology to the smaller subunit and high homology in the domain found in all cyclases.

Correspondence address: D. Koesling, Institut für Pharmakologie Freie Universität Berlin, Thielallee 67-73, D-1000 Berlin 33, FRG

2. MATERIALS AND METHODS

For the purification of soluble guanylyl cyclase from bovine lung, we used an immunoaffinity method with immobilized peptide antibodies directed against a synthetic C-terminal peptide of the 70 kDa subunit of the enzyme [15]. With this method, we obtained an apparently homogenous enzyme which was stimulated up to 250-fold by sodium nitroprusside and revealed a larger subunit with an M_r of about 73 kDa on SDS gels. About 0.1 mg of the purified enzyme was separated on a 10% SDS gel. The band corresponding to the 73 kDa subunit of soluble guanylyl cyclase was excised from the gel, cut into small pieces (about 1 × 1 mm) and washed in water with frequent changes over 16 h. The washed gel pieces were then immersed in 0.2 ml of 0.1 M ammonium hydrogen carbonate, pH 8.5, containing trypsin at an enzyme-to-protein ratio of about 1:5. Following incubation at 37°C for 8 h, the generated protein fragments were eluted from the gel by shaking the pieces twice for 3 h with an equal volume of 0.1% trifluoroacetic acid in water (v/v). Residual water was extracted from the gel matrix by treatment with acetonitrile. The concentrated eluates were extracted twice with a 1:4 mixture of isoamylalcohol/heptane to remove traces of SDS [16]. Separation of the tryptic peptides was performed by reversed-phase HPLC on a Vydac 218TP5 column (1.6 × 250 mm). The peptides were sequenced in a gas phase sequencer according to [17].

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

A cDNA library from bovine brain constructed in λgt10 and a cDNA library from bovine adrenal medulla constructed in λZAP by Clontech were bought from Genofit, Heidelberg. The procedures for screening were as described by Maniatis et al. [18]. The oligonucleotide probe

```
  C  A  A  T  G  G  G
AT TC AA TA TC TC AA TG
  T  G  G  C  A  A  A
```

labelled with ^{32}P at the 5′-end was used for hybridization at 37°C. Washing was carried out at 37°C in 6 × SSC (1 × SSC = 0.15 M NaCl/15 mM sodium citrate, pH 7.0) with 0.1% SDS (w/v). The insert of clone 27 (bp 1–bp 2046) was radioactively labelled by random oligonucleotide priming and extension of double stranded DNA [19]. Using radiolabelled cDNA as a probe for screening libraries, hybridization was performed at 65°C. Washing was carried out at 65°C in 0.1 × SSC and 0.1% SDS (w/v).

For Northern blots, total RNA was isolated as described by Turpen and Griffith [20]. Poly(A)$^+$ RNA was purified by oligo-(dT)-cellulose column chromatography. 14 μg of poly(A)$^+$ RNA were glyoxylated, separated on agarose gels and blotted on GeneScreen nylon membranes. Hybridization at 65°C was carried out with a 1.8 kb cDNA insert from clone 27, labelled as described above. Washing of the Northern blots was performed according to the GeneScreen instruction manual. DNA was sequenced with the dideoxy chain termination method [21,22].

3. RESULTS AND DISCUSSION

Seven internal peptides (**1** = IVVDQAIAAGVPVEII, **2** = ILYETEVEVS, **3** = TFPFHFML, **4** = PHFDEYFEILT, **5** = DVVLIG, **6** = MMELSHEVVSPHGEP, **7** = DXPGFVFTP; Fig. 1) of the isolated 73 kDa subunit were determined by sequence analysis, providing partial sequence information throughout the protein. In contrast to the N-terminus of the 70 kDa subunit, the N-terminus of the 73 kDa subunit appears to be blocked as several attempts to obtain an N-terminal sequence failed; the nature of this block remains unknown.

A cDNA library from bovine brain was screened by hybridization with radiolabelled oligonucleotides representing backtranslations of a part of the amino acid sequence of peptide **4** (HFDEYFEI). One positive clone, designated 27, with an insert of 1836 bp was isolated; the insert was subcloned and sequenced. The sequence contained an open reading frame from an ATG in position 291 until the end of the clone. In the deduced amino acid sequence of clone 27, the complete sequence of peptide **4** (bp 1258–bp 1281), which had been used for screening, and the sequences of 4 other peptides derived from protein sequencing were detected. Since clone 27 did not contain the C-terminus, clone 27 was used as a probe to screen a cDNA library from bovine adrenal medulla. One clone, designated K01, with an insert of 2985 bp was isolated; the insert was subcloned and sequenced. The 5′-end of clone K01 overlapped with clone 27 from bp 275 of clone 27 and, thereby, contained the putative initiation codon and a termination codon at base 2386 (see Fig. 1). The following non-coding region of 898 bp at the 3′-end of clone K01 did not contain a polyadenylation signal. The open reading frame encoded a polypeptide with a predicted molecular mass of 77 500, which is in reasonable agreement with the M_r of about 73 kDa observed on SDS gels. All 7 peptides derived from protein sequencing were detected in the deduced amino acid sequence. The potential initiation codon is preceded by the sequence CACC, which fullfils Kozak's rules as a consensus sequence for eukaryotic initiation sites [23], and by a GC-rich sequence containing stop codons in all frames. The next ATG towards the 3′-end is at position 1165 and, therefore, downstream of peptides **1** and **2** and within peptide **3** known to be part of the sequence. A hydropathy plot of the predicted polypeptide did not show any large hydrophobic or hydrophilic regions (not shown).

Blot hybridization analysis of bovine brain poly(A)$^+$ RNA showed a major band corresponding to a RNA of about 3.8 kb and a minor band with a size of about 2.9 kb (not shown). The major band probably represents the mRNA coding for the larger subunit of soluble guanylyl cyclase. The minor band may indicate the existence of an mRNA coding for an isoenzyme.

A comparison of the deduced amino acid sequences from the 73 kDa subunit of soluble guanylyl cyclase with the 70 kDa one revealed strong similarities between both subunits (Fig. 2). The two subunits of soluble guanylyl cyclase, lined up as in Fig. 2, share 201 (about 32%) identical amino acids. The relative amount of identical amino acids varies in different homologous regions of the sequences. The first region (318 amino acids) of the alignment shows a rather low similarity with 20% identical amino acids over this region, but contains 4 conserved cysteine residues. In the following region of 54 amino acids, the highest rate of identical amino acids (70%) is found. The similarity continues with about 40% identical residues along the next 231 amino acids. In contrast, in the C-terminal part (46 amino acids), there are almost no identities.

Like the 70 kDa subunit of soluble guanylyl cyclase, the 73 kDa subunit shares homologies in a region of about 250 amino acids with the plasma membrane-bound, atrial natriuretic peptide (ANP)-stimulated guanylyl cyclases of type A and B from rat brain, the plasma membrane-bound guanylyl cyclase of sperm from the sea urchin, *Strongylocentrotus purpuratus,* as well as with the two hydrophilic domains of the adenylyl cyclase from bovine brain (see Fig. 2). The alignment reveals 6 highly conserved clusters of amino acids among the guanylyl cyclases. The similarity with the adenylyl cyclase is less pronounced. In comparison with the protein sequence data library of the National Biomedical Research Foundation (NBRF) and with a translation of the European Molecular Biology Laboratory (EMBL) nucleotide sequence database, the sequence of the 73 kDa subunit shows no other significant homologies.

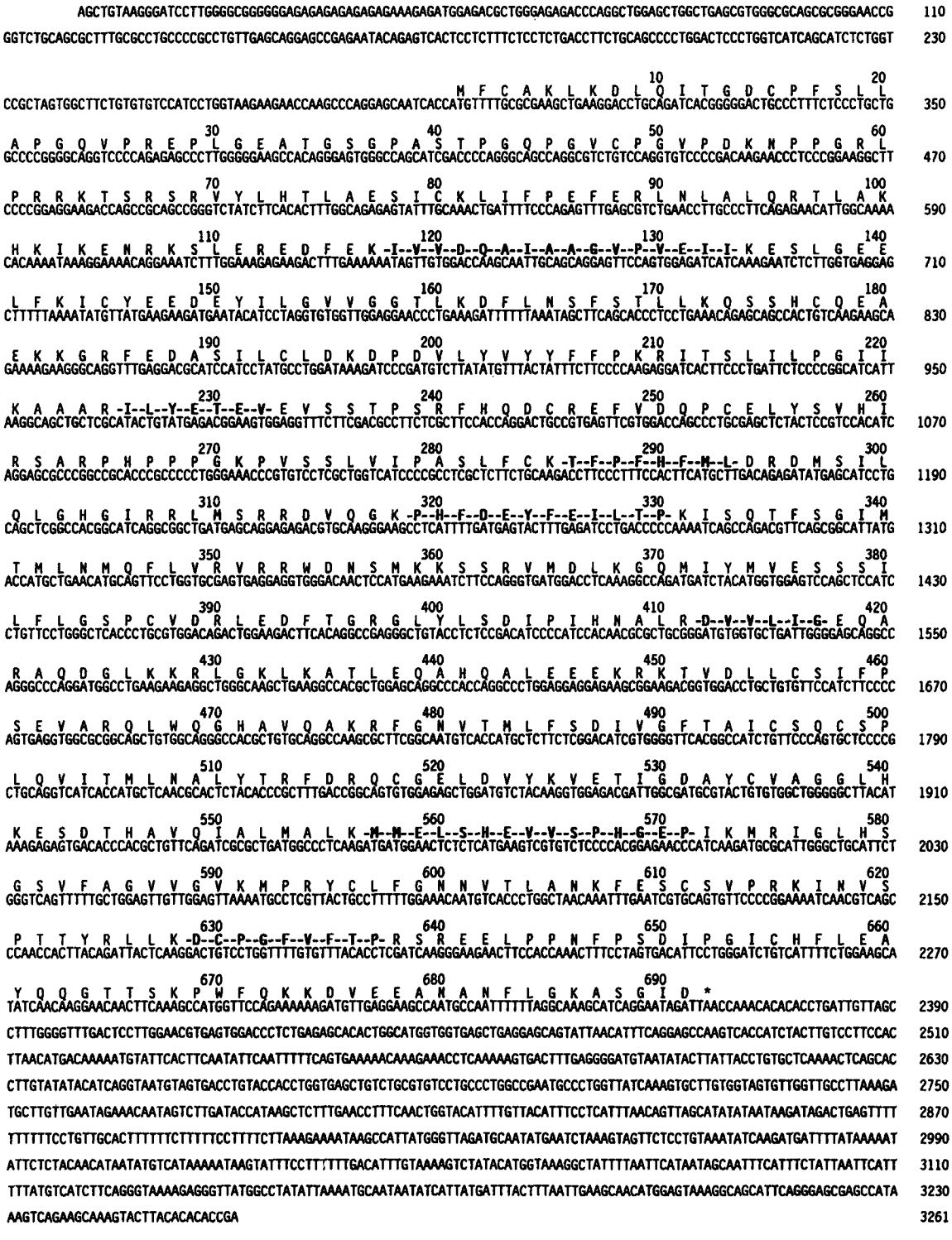

Fig. 1. Complete nucleotide and deduced amino acid sequences of the 73 kDa subunit of soluble guanylyl cyclase. The peptide sequences obtained from protein sequencing are indicated by bold letters and connected by hyphens.

The sequence of the larger subunit of soluble guanylyl cyclase shows strong homologies to the sequence of the 70 kDa subunit including the region possessed by all cyclases. The cyclases show very different mechanisms of regulation by coupling to G-proteins, by direct hormone binding and by NO-containing compounds for respectively adenylyl cyclase, plasma membrane-bound and soluble guanylyl cyclases, but share essentially the same catalytic functions. We presume that the C-terminal, homologous

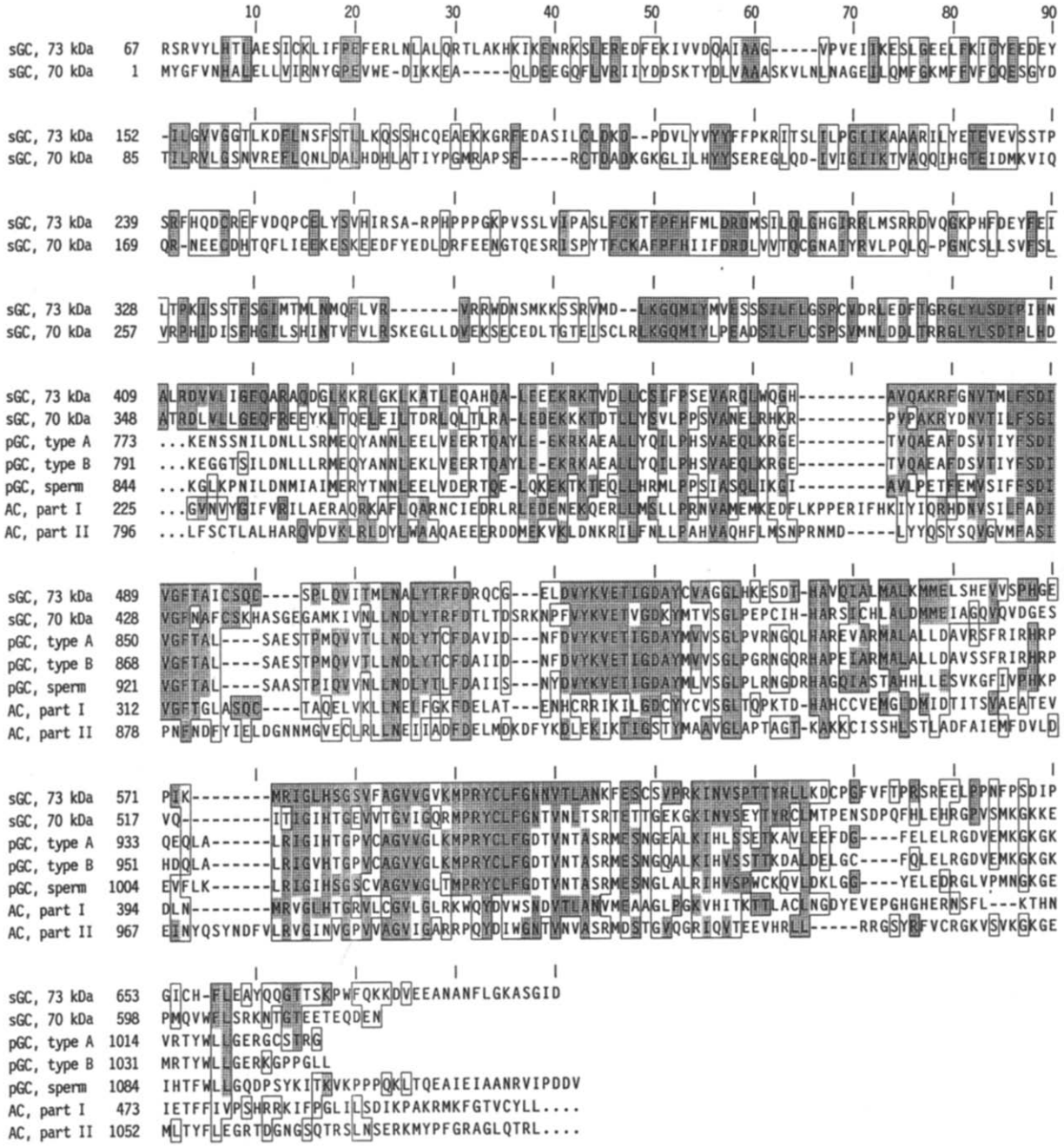

Fig. 2. Alignment of amino acid sequences of the subunits of bovine soluble guanylyl cyclase and of 3 plasma membrane-bound guanylyl cyclases and the two hydrophilic domains of a mammalian adenylyl cyclase. The sequences shown are the larger subunit of bovine soluble guanylyl cyclase (sGC, 73 kDa, aa 67–691, total aa 691) and the smaller subunit (sGC, 70 kDa, aa 1–619, total aa 619 [10], the ANP-regulated guanylyl cyclase type A (pGC, type A, aa 773–1029, total aa 1029) and type B (pGC, type B, aa 791–1047, total aa 1047) both from rat brain [5,8], the resact-regulated guanylyl cyclase (pGC, sperm, aa 844–1125, total aa 1125) from the sperm of sea urchin, *Strongylocentrotus purpuratus* [6], and two selected regions of the adenylyl cyclase (AC, part I, aa 225–510, and AC, part II, aa 796–1089, total aa 1134) from bovine brain [12]. Amino acid sequences were compared with the 73 kDa subunit of the soluble guanylyl cyclase using the Diagon program by Staden [24], and alignments were optimized. Conservative substitutions are boxed and identical amino acids are shadowed. Conservative substitutions were selected according to the amino acid score matrix MDM78 of Dayhoff [25] except that W was not substituted by R, but by Y and F.

region, common to all sequences, represents a domain involved in the catalytic function of the cyclases. It should be pointed out that both the soluble guanylyl cyclase, occurring as a dimer, and the adenylyl cyclase contain two of these homologous domains. It has to be established whether two of these regions are required for catalytic function of all cyclases. If so, a dimerization of the membrane-bound guanylyl cyclases has to be postulated.

Acknowledgements: We thank Dr Toby Gibson for helpful discussion as well as for help in aligning the sequences and Dr B. Sproat and P. Neuner for synthesizing the oligonucleotides. This work was supported by the Deutsche Forschungsgemeinschaft, Bo 888/1-1, and by the Fonds der Chemischen Industrie.

REFERENCES

[1] Mittal, C.K. and Murad, F. (1982) in: Handbook of Experimental Pharmacology, vol. 58/I (Nathanson, J.A. and Kebabian, J.W. eds) pp. 225–260, Springer-Verlag, Berlin.
[2] Waldman, S.A. and Murad, F. (1987) Pharmacol. Rev. 39, 163–196.
[3] Lincoln, T.M. and Corbin, J.D. (1983) Adv. Cycl. Nucl. Res. 15, 139–192.
[4] Fesenko, E.E., Kolesnikov, S.S. and Lyubarsky, A.L. (1985) Nature 313, 310–313.
[5] Chinkers, M., Garbers, D.L., Chang, M.-S., Lowe, D.G., Chin, H., Goeddel, D.V. and Schulz, S. (1989) Nature 338, 78–83.
[6] Singh, S., Lowe, D.G., Thorpe, D.S., Rodriguez, H., Kuang, W.-J., Dangott, L.J., Chinkers, M., Goeddel, D.V. and Garbers, D.L. (1988) Nature 334, 708–712.
[7] Lowe, D.G., Chang, M.-S., Hellmiss, R., Chen, E., Singh, S., Garbers, D.L. and Goeddel, D.V. (1989) EMBO J. 8, 1377–1384.
[8] Schulz, S., Singh, S., Bellet, R.A., Singh, G., Tubb, D.J., Chin, H. and Garbers, D.L. (1989) Cell 58, 1155–1162.
[9] Chang, M.-S., Lowe, D.G., Lewis, M., Hellmiss, R., Chen, E. and Goeddel, D.V. (1989) Nature 341, 68–72.
[10] Koesling, D., Herz, J., Gausepohl, H., Niroomand, F., Hinsch, K.-D., Mülsch, H., Böhme, E., Schultz, G. and Frank, R. (1988) FEBS Lett. 239, 29–34.
[11] Nakane, M., Saheki, S., Kuno, T., Ishii, K. and Murad, F. (1988) Biochem. Biophys. Res. Commun. 157, 1139–1147.
[12] Krupinski, J., Coussen, F., Bakalyar, H.A., Tang, W.-J., Feinstein, P.G., Orth, K., Slaughter, C., Reed, R.R. and Gilman, A.G. (1989) Science 244, 1558–1564.
[13] Gerzer, R., Hofmann, F. and Schultz, G. (1981) Eur. J. Biochem. 116, 479–486.
[14] Kamisaki, Y., Saheki, S., Nakane, M., Palmieri, J., Kuno, T., Chang, B., Waldman, S.A. and Murad, F. (1986) J. Biol. Chem. 261, 7236–7241.
[15] Humbert, P., Niroomand, F., Fischer, G., Mayer, B., Koesling, D., Hinsch, K.-H., Gausepohl, H., Frank, R., Schultz, G. and Böhme, E. (1990) Eur. J. Biochem., in press.
[16] Bosserhoff, A., Wallach, J. and Frank, R. (1989) J. Chromatogr., 473, 71–77.
[17] Gausepohl, H., Trosin, M. and Frank, R. (1986) in: Advanced Methods in Protein Microsequence Analysis (Wittmann-Liebold, M. ed.), pp. 149–160, Springer-Verlag, Berlin.
[18] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Publications, Cold Spring Harbor, NY.
[19] Feinberg, A.P. and Vogelstein, B. (1983) Anal. Biochem. 132, 6–13.
[20] Turpen, T.H. and Griffith, O.M. (1986) Biotechniques 4, 14–15.
[21] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[22] Messing, J. (1983) Methods Enzymol. 101, 20–78.
[23] Kozak, M. (1984) Nucleic Acids Res. 12, 857–873.
[24] Staden, R. (1982) Nucleic Acids Res. 10, 2951–2961.
[25] Dayhoff, M.O. (1969) Atlas of Protein Sequence and Structure, National Biochemical Research Foundation, Silver Springs, MD.