Volume 298, number 2,3, 137–141 FEBS 10726 February 1992

Sequence motif in control regions of the H^+/K^+ ATPase α and β subunit genes recognized by gastric specific nuclear protein(s)

Shigehiko Tamura[a], Ko-Ichi Oshiman[a], Tsuyoshi Nishi[a], Masaharu Mori[b], Masatomo Maeda[a] and Masamitsu Futai[a]

[a]*Department of Organic Chemistry and Biochemistry, Institute of Scientific and Industrial Research, Osaka University, Ibaraki, Osaka 567, Japan* and [b]*Department of Pathology, Okayama University Medical School, Okayama 700, Japan*

Received 16 December 1991

A nuclear protein(s) from rat or pig stomach recognized a conserved sequence in the 5′-upstream regions of the rat and human H^+/K^+-ATPase α subunit genes. A gel retardation assay suggested that part of the binding site was located in the **TAATCAGCTG** sequence. No nuclear proteins capable of the binding could be detected in other tissues of rat (liver, brain, kidney, spleen and lung) or pig liver. The sequence motif (**GATAGC**) located 5′-upstream of the β-subunit gene also seemed to be recognized by the same protein, because the binding of nuclear protein to the sequence motifs in the α and β subunits was mutually competitive. Considering the sense-strand sequence of the binding motif in the α-subunit gene, we conclude that **(G/C)PuPu(G/C)NGAT(A/T)PuPy** is a core sequence motif for the gastric specific DNA binding protein (PCSF, parietal cell specific factor).

H^+/K^+-ATPase; Gastric specific transcription; Gastric proton pump; Parietal cell; DNA binding protein

1. INTRODUCTION

Gastric parietal cells are well known to be highly differentiated eukaryotic cells. These cells have a characteristic morphology and secrete HCl into the gastric lumen [1], synthesizing specific enzymes essential for the secretion. The genes for the gastric proton pump (H^+/K^+-ATPase) α and β subunits have been shown to be transcribed [2–6], and translated [1,3,5] only in the gastric mucosa (parietal cell). Recently, we reported that the exon/intron organizations of the genes for the H^+/K^+-ATPase α and β subunits are highly similar to those of the corresponding subunits of Na^+/K^+-ATPase [6–8]. However, the nucleotide sequences in the control regions of the genes for the H^+/K^+-ATPase α [6,7] and β [8] subunits have no apparent similarity with those of Na^+/K^+-ATPase subunits [9–14]. Similarity is evident in the 5′-upstream regions of the rat and human H^+/K^+-ATPase α-subunit genes [6], suggesting that these regions may be important for the transcriptional regulation of the α-subunit gene.

In this study we found that gastric mucosal nuclear protein(s) recognized a sequence motif in the 5′-upstream regions of both the rat and human genes for the α subunit of H^+/K^+-ATPase. This gastric specific protein also bound to the motif in the rat β-subunit gene. The consensus core sequence motif for the α- and β-subunit genes is presented.

Abbreviations: HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; EDTA, ethylenediamine-tetraacetic acid

Correspondence address: S. Tamura, Department of Organic Chemistry and Biochemistry, Institute of Scientific and Industrial Research, Osaka University, Ibaraki, Osaka 567, Japan. Fax: (81) (6) 875 5724.

2. MATERIALS AND METHODS

2.1. *Preparation of nuclear proteins from various tissues*

Nuclear proteins were prepared at 4°C as described previously [15] with minor modifications. Minced tissue (50 g of pig gastric mucosa, 15 g of pig liver or 10 g of various rat tissues) was brought to a volume of 100 ml with homogenization buffer (10 mM HEPES-KOH pH 7.6, 15 mM KCl, 0.15 mM spermine, 0.5 mM spermidine, 1 mM EDTA, 2.2 M sucrose, 5% glycerol, 0.5 mM dithiothreitol, 0.5 mM phenylmethanesulfonyl fluoride, 1 mg/ml pepstatin), and homogenized at 900 rpm in a motor-driven Potter Elvehjem homogenizer. The homogenate was filtered through cheese cloth and 50 ml aliquots were layered over 10 ml of cushion buffer (the same buffer as for homogenization but with 2 M sucrose and 10% glycerol were used), and centrifuged at 130,000 × *g* for 1 h in an SW28 rotor (Hitachi ultracentrifuge). The nuclear pellet was resuspended in 60 ml of lysis buffer (10 mM HEPES-KOH pH 7.6 containing 100 mM KCl, 0.1 mM EDTA, 10% glycerol, 3 mM $MgCl_2$, 1 mM dithiothreitol, 0.1 mM phenylmethanesulfonyl fluoride, 1 mg/ml pepstatin). The suspension was diluted to 10 A_{260} U/ml with the lysis buffer, adjusted to 0.55 M KCl by adding 3 M KCl, shaken gently for 30 min, and centrifuged at 100,000 × *g* for 1 h. Solid $(NH_4)_2SO_4$ (0.3 g/ml) was slowly added to the supernatant and the solution was incubated for 60 min. The precipitate was collected by centrifugation (100,000 × *g*, 1 h), dissolved in 1 ml of 25 mM HEPES-KOH pH 7.6, 0.1 mM EDTA, 40 mM KCl, 10% glycerol and 1 mM dithiothreitol, and dialyzed against 500 vols. of the same buffer for 4 h. The precipitate formed during dialysis was removed by centrifugation (10,000 × *g*, 5 min), and the nuclear protein fraction (2 mg/ml) was stored at −70°C. Protein was measured with a Bio-Rad protein assay kit [16] using bovine serum albumin as a standard.

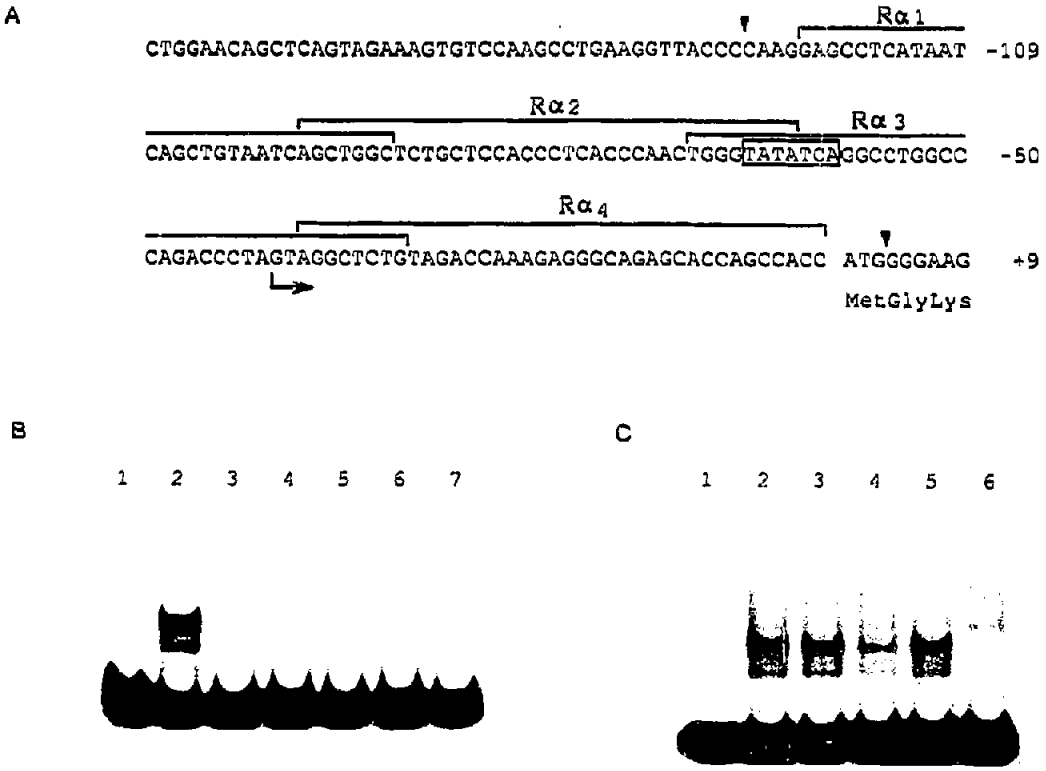

Fig. 1. Detection of nuclear proteins recognizing the 5'-upstream region of the rat gastric H/K-ATPase α-subunit gene. A. The 5'-upstream region (designated as RA, −124 to +3 bp) between the arrow heads of the rat gastric H/K-ATPase α-subunit gene [6] was divided into four segments: Rα1 (−120 to −91 bp), Rα2 (−97 to −62 bp), Rα3 (−69 to −81 bp) and Rα4 (−38 to −1 bp). The positions of the potential TATA-box [20] and transcription initiation site [21] are shown by box and arrow, respectively. Nucleotides are numbered on the right of each line from the first letter of the initiation codon. B. Nuclear extracts (0.16 μg) from rat tissues were incubated with radio-labeled segment RA (0.1 pmol). Complex formation was determined by gel retardation assay (lane 1, without protein; lanes 2–7, with nuclear protein). Stomach (lane 2), liver (lane 3), kidney (lane 4), spleen (lane 5), brain (lane 6) and lung (lane 7). C. Nuclear extract from stomach (0.16 μg) was incubated with the radio-labeled DNA segment RA (0.1 pmol) (lane 1, without protein; lanes 2–6, with nuclear protein). Rα4 (lane 3), Rα3 (lane 4), Rα2 (lane 5) or Rα1 (lane 6) was added as a competitor.

2.2. *Gel retardation assay*

Gel retardation assays [17] were performed in 15 μl of 25 mM HEPES-KOH (pH 7.6) containing 0.1 pmol of P-end-labeled DNA probe (3,000–5,000 cpm), 2 μg of poly (dI-dC), 34 mM KCl, 5 mM MgCl, and 5% glycerol. Binding was started by adding nuclear protein (0.16–0.28 μg). Mixtures were incubated for 30 min on ice, and the DNA–protein complexes formed were separated from free DNA probe at 4°C by 8% polyacrylamide gel electrophoresis in 89 mM Tris-borate (pH 8.0) containing 2.5 mM EDTA. Electrophoresis buffer was circulated. Gels were dried and autoradiographed. The 5' ends of probes were labeled with [α-P]dCTP using the Klenow fragment. Free [α-P]dCTP was removed by polyacrylamide gel (5%) electrophoresis. Synthetic double-stranded oligonucleotides (synthesized in an Applied Biosystems Synthesizer 381A) were used as competitors.

2.2. *Chemicals*

Restriction enzymes, T4 DNA ligase, the Klenow fragment of *E. coli* DNA polymerase I, DNase I and Taq DNA polymerase were purchased from Takara Shuzo Co., Kyoto, Japan. [α-P]dCTP (3,000 Ci/mmol) was from the Radiochemical Center, Amersham Corp. All other chemicals used were of the highest grade commercially available.

3. RESULTS

3.1. *The 5'-upstream region of the H/K-ATPase α-subunit gene*

We cloned the region between −1,149 and +136 bp of the human H/K-ATPase α-subunit gene [7] from gastric mucosa and liver of the same person (autopsy materials) using the polymerase chain reaction [18]. Parietal cells are the most abundant cell type (more than 50%) in gastric mucosa [1,3,6]. The cloned sequences from the two tissues were identical and exactly the same as reported [7], except that one base (T) in an *Alu* sequence (position −601) was substituted by C, possibly due to personal deviation. These results suggested that the 5'-upstream sequence of the α-subunit gene is identical in different tissues, and that a specific *trans*-acting protein(s) participates in expression of the gastric gene.

3.2. *Binding of nuclear proteins to the 5'-upstream region of the α-subunit gene*

The above observation prompted us to try to detect a specific *trans*-acting protein(s) in gastric nuclei. The DNA segment (RA) immediately flanking the initiation codon of the rat α-subunit gene [6] (Fig. 1A) was incubated with nuclear proteins prepared from various rat tissues. This segment carried a TATA-like sequence and had homology with that of the human α-subunit gene.

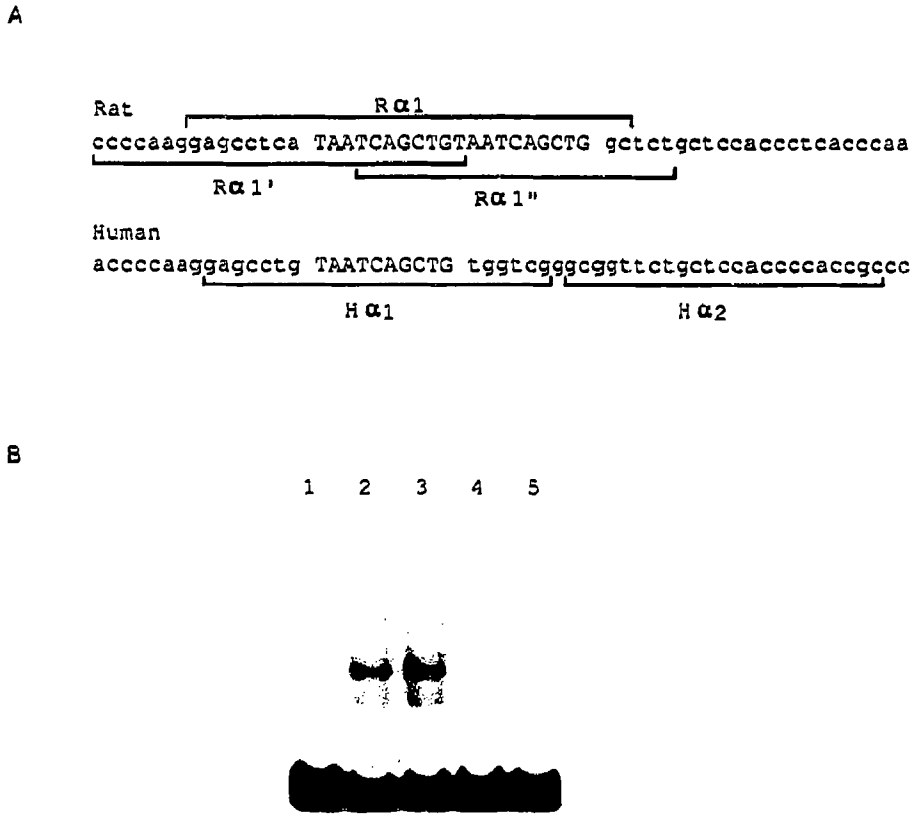

Fig. 2. Determination of protein binding motif in the α subunit. A. The 5'-upstream regions of rat (upper) and human (lower) α-subunit genes are shown. Synthetic double-stranded DNAs used as competitors were: Rα1 (−120 to −91 bp) for the rat α-subunit gene [6] and Hα1 (−116 to −94 bp) and Hα2 (−92 to −70 bp) for the human gene [7]. The positions of Rα1′ (−127 to −102 bp) and Rα1″ (−109 to −88 bp) are also indicated (see Fig. 4). **TAATCAGCTG** sequences are shown by capital letters. B. Nuclear extract from rat stomach (0.28 μg) and radio-labeled segment HA (−127 to +3 bp, ref. [7]) (0.1 pmol) were subjected to gel retardation assay (lane 1, without protein; lanes 2–5, with nuclear protein). Hα2 (lane 3), Hα1 (lane 4) or Rα1 (lane 5) was added as a competitor.

Shifted bands due to complexes between the gastric nuclear protein(s) and the DNA segment were detected by gel electrophoresis (Fig. 1B, lanes 1 and 2). Nuclear proteins that had been heated at 55°C for 3 min did not form these complexes with segment RA. Proteins from liver, brain, kidney, spleen and lung also did not form complexes (Fig. 1B, lanes 3–7). We also found that only the nuclear protein(s) from stomach bound to the corresponding human segment HA (*Bst*EII–*Nco*I fragment, positions −127 to +3 bp, Ref. [7]) (not shown).

3.3. *Sequence motif for protein binding*

To determine the protein binding region more precisely, we synthesized double-stranded oligonucleotides (Rα1, Rα2, Rα3 and Rα4) (Fig. 1A) and tested these as competitors in the binding assay. Only Rα1 prevented the formation of complexes between the nuclear protein and radio-labeled segment RA (Fig. 1C). Rα1 contained tandem **TAATCAGCTG** sequences, whereas a single copy of the sequence was found in the corresponding region of the human gene (Fig. 2A). Radio-labeled HA formed complexes with gastric nuclear protein as shown by a gel retardation assay (Fig. 2B). The formation of these complexes between DNA (HA) and nuclear protein was inhibited by unlabeled Hα1 (the human DNA segment corresponding to Rα1 of rat DNA) as well as Rα1. The human Hα2 segment did not affect complex formation. A single stranded DNA fragment (sense or anti-sense strand) of Hα1 did not bind nuclear proteins (data not shown). These results suggest that the **TAATCAGCTG** sequence could be part of the binding site of gastric nuclear protein.

3.4. *Comparison of sequence motifs in the α- and β-subunit genes*

We suggested recently that the **GATAGC** motif in the 5'-upstream region of the rat β-subunit gene is recognized by a specific nuclear protein from rat stomach [8]. Since the α- and β-subunit genes are both specifically transcribed in parietal cells, it was of interest to examine whether the sequence motifs for the α- and β-subunit genes are recognized by the same protein. The protein binding with radioactive segment RA (carrying the **TAATCAGCTG** sequence) became undetectable in the presence of the segments carrying the **GATAGC** sequence (Rβ2, Rβ3 and Rβ4, Ref. [8]) or the segment

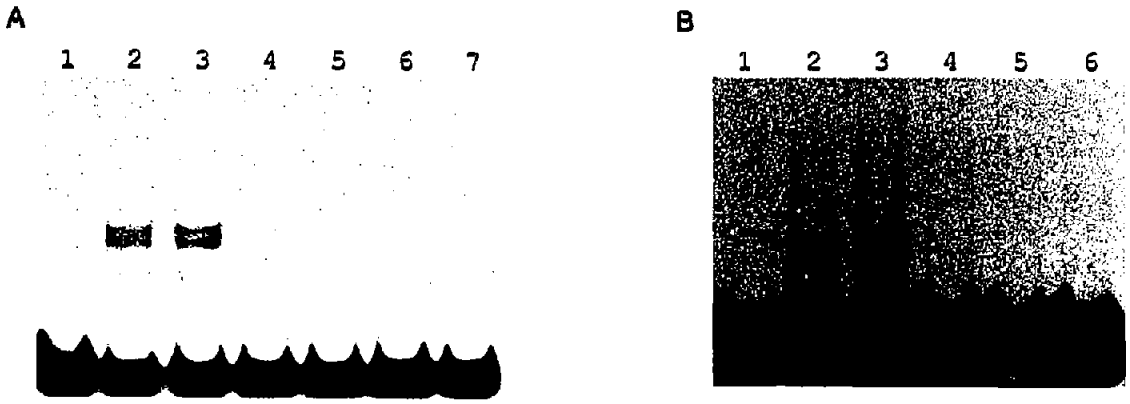

Fig. 3. Mutual competition of protein binding to the sequence motifs for the α- and β-subunit genes. Nuclear extract from rat stomach (0.16 μg) was incubated with the radio-labeled DNA segments (0.1 pmol) of the α- and β-subunit genes. A. Segment RA of the rat α-subunit gene (Fig. 1) was used as a probe (lane 1, without protein; lanes 2–7, with nuclear protein). Rβ5 (lane 3), Rβ4 (lane 4), Rβ3 (lane 5), Rβ2 (lane 6) or Rα1 (lane 7) was added as a competitor. B. The segment (−185 to +10 bp) of the rat β subunit gene [8] was used as a probe (lane 1; without protein, lanes 2–6; with nuclear protein). Rβ1 (lane 3), Rβ2 (lane 4), Rα1 (lane 5) or Hα1 (lane 6) was added as a competitor. See Ref. [8] and Fig. 2A.

carrying the **TAATCAGCTG** sequence (Rα1) (Fig. 3A). However, segment Rβ5 without the **GATAGC** sequence [8] was not a competitor. Furthermore, binding of the protein to the 5′-upstream region of the β-subunit gene carrying three **GATAGC** sequences (−185 to +10 bp) [8] was abolished by Rα1 and Hα1 as well as Rβ2, but not by Rβ1 without the **GATAGC** sequence [8] (Fig. 3B). These results strongly suggested that the same or a closely related protein(s) binds to the sequence motifs in the α- and β-subunit genes.

Comparison of the sense-strand of the sequence motif in the α-subunit gene with that in the β subunit showed that **(G/C)PuPu(G/C)NGAT(A/T)PuPy** (Pu, A or G; Py, C or T; N, A, C or T) is the common core sequence motif recognized by the gastric specific nuclear protein(s) (Fig. 4). The nuclear protein may bind to both the sequence motifs tandemly repeated in the α-subunit gene (Fig. 2A): Rα1′ was also a competitor in gel retardation assay (not shown), and the sequence Rα″ is essentially the same as that of Hα1. A nuclear protein(s) with essentially the same properties as that from rat stomach was obtained from pig gastric mucosa (but not liver) (data not shown).

4. DISCUSSION

A specific protein(s) from rat stomach recognized the conserved sequence located upstream of the human and rat α-subunit genes. The same nuclear protein also bound to the upstream region of the rat β-subunit gene. A similar DNA binding protein(s) was found in a nuclear extract of pig gastric mucosa. These results suggest that in mammals, gastric acid secretion is governed essentially by the same transcriptional regulatory mechanism. The common core sequence motif for the nuclear protein(s) was suggested to be **(G/C)PuPu(G/C)NGAT(A/T)PuPy.**

A major DNA-binding protein (GF-1) in cells of erythroid lineage is known to recognize the **(A/T)GATA(A/G)** core sequence [19]. The motif for gastric protein includes this core sequence, suggesting that the gastric specific DNA-binding protein and GF-1 may be members of a family of specific transcriptional regulatory factors. The motifs are located about 30 bp upstream of the conserved TATA-like box, like other known regulatory protein binding sites [20]. The transcription initiation sites were demonstrated to be 25 bp and 30 bp downstream of the TATA-like boxes for

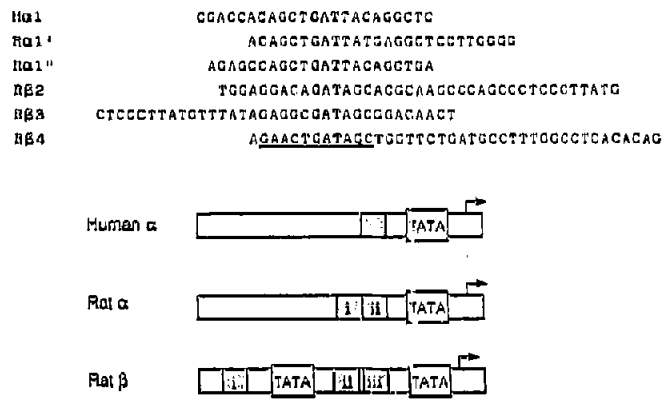

Fig. 4. Sequence comparison of the motifs in the α- and β-subunit genes. A. The sequence motifs (sense-strand) in the α-subunit genes were compared with those in the β subunit. The DNA segments shown in Fig. 2 and Ref. [8] were aligned. The sequence corresponding to the common core sequence motif is underlined. B. The positions of the motifs in the human α-, rat α- and rat β-subunit genes are shown schematically as dotted boxes; the motifs in Rα1′ and Rα1″, and Rβ2, Rβ3 and Rβ4 shown above corresponded to the boxes i and ii in Ratα, and i, ii and iii in Ratβ, respectively. Transcriptional start sites [21,24] were shown by arrows. Only the major start site of the rat β-subunit gene is shown.

the human and rat α-subunit genes, respectively [21]. The motif identified in this study may be a binding site for a positive transcriptional regulator that functions specifically in parietal cells, because the protein was found in stomach but not in the other tissues examined. We named the protein PCSF (parietal cell specific factor) (Fig. 4). Three shifted bands for protein–DNA complexes were detected by gel electrophoresis. This may be due to the presence of DNA binding proteins that are phosphorylated [22] and/or associated with other proteins [23].

Acknowledgements: This work was supported by grants from the Ministry of Education, Science and Culture of Japan, a Human Frontier Science Program Research Grant, and a grant from the Naito Foundation. We thank Dr. K. Sekimizu (Tokyo University) for helpful discussions.

REFERENCES

[1] Saccomani, G., Helander, H.F., Crago, S., Chang, H.H., Dailey, D.W. and Sachs, G. (1979) J. Cell. Biol. 83, 271–283.
[2] Shull, G.E. (1990) J. Biol. Chem. 265, 12123–12126.
[3] Toh, B.-H., Gleeson, P.A., Simpson, R.J., Moritz, R.L., Callaghan, J.M., Goldkorn, I., Jones, C.M., Martinelli, T.M., Mu, F.-T., Humphris, D.C., Pettitt, J.M., Mori, Y., Masuda, T., Sobieszczuk, P., Weinstock, J., Mantamadiotis, T. and Baldwin, G.S. (1990) Proc. Natl. Acad. Sci. USA 87, 6418–6422.
[4] Reuben, M.A., Lasater, L.S. and Sachs, G. (1990) Proc. Natl. Acad. Sci. USA 87, 6767–6771.
[5] Canfield, V.A., Okamoto, C.T., Chow, D., Dorfman, J., Gros, P., Forte, J.G. and Levenson, R. (1990) J. Biol. Chem. 265, 19878–19884.
[6] Oshiman, K., Motojima, K., Mahmood, S., Shimada, A., Tamura, S., Maeda, M. and Futai, M. (1991) FEBS Lett. 281, 250–254.
[7] Maeda, M., Oshiman, K., Tamura, S. and Futai, M. (1990) J. Biol. Chem. 265, 9027–9032.
[8] Maeda, M., Oshiman, K., Tamura, S., Kaya, S., Mahmood, S., Reuben, M.A., Lasater, L.S., Sachs, G. and Futai, M. (1991) J. Biol. Chem. 266, 21584–21588.
[9] Kano, I., Nagai, F., Satoh, K., Ushiyama, K., Nakao, T. and Kano, K. (1989) FEBS Lett. 250, 91–98.
[10] Sverdlov, E.D., Bessarab, D.A., Malyshev, I.V., Petrukhin, K.E., Smirnov, Y.V., Ushkaryov, Y.A., Monastyrskaya, G.S., Broude, N.E. and Modyanov, N.N. (1989) FEBS Hett. 244, 481–483.
[11] Shull, M.M., Pugh, D.G. and Lingrel, J.B. (1989) J. Biol. Chem. 264, 17532–17543.
[12] Kawakami, K., Yagawa, Y. and Nagano, K. (1990) Gene 91, 267–270.
[13] Ovchinnikov, Y.A., Monastyrskaya, G.S., Broude, N.E., Ushkaryov, Y.A., Melkov, A.M., Smirnov, Y.V., Malyshev, I.V., Allikmets, R.L., Kostina, M.B., Dulubova, I.E., Kiyatkin, N.I., Grishin, A.V., Modyanov, N.N. and Sverdlov, E.D. (1988) FEBS Lett. 233, 87–94.
[14] Lane, L.K., Shull, M.M., Whitmer, K.R. and Lingrel, J.B. (1988) Genomics 5, 445–453.
[15] Gorski, K., Carneiro, M. and Schibler, U. (1986) Cell 47, 767–776.
[16] Bradford, M.M. (1976) Anal. Biochem. 72, 248–254.
[17] Fried, M. and Crothers, D.M. (1981) Nucleic Acids Res. 9, 6505–6525.
[18] Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) Science 239, 487–491.
[19] Tsai, S.-F., Martin, D.I.K., Zon, L.I., D'Andrea, A.D., Wong, G.G. and Orkin, S.H. (1989) Nature 339, 446–451.
[20] Struhl, K. (1989) Annu. Rev. Biochem. 58, 1051–1077.
[21] Newman, P.R., Greeb, J., Keeton, T.P., Reyes, A.A. and Shull, G.E. (1990) DNA and Cell Biol. 9, 749–762.
[22] Jackson, S.P., MacDonald, J.J., Lees-Miller, S. and Tjian, R. (1990) Cell 63, 155–165.
[23] Ptashne, M. and Gann, A.A.F. (1990) Nature 346, 329–331.
[24] Newman, P.R. and Shull, G.E. (1991) Genomics 11, 252–262.