Vol. 164, No. 1, 1989 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
October 16, 1989 Pages 176-184

Isolation and Characterization of a Molecular cDNA Clone of a Human mRNA from Interferon-treated Cells Encoding Nucleolar Protein B23, Numatrin*,[1]

Xuxian T. Zhang, Daniel C. Thomis and Charles E. Samuel

Section of Biochemistry and Molecular Biology
Department of Biological Sciences
University of California, Santa Barbara, CA 93106

Received August 21, 1989

SUMMARY: A cDNA clone encoding human nucleolar phosphoprotein B23, numatrin, was isolated from a library prepared with mRNA from human U cells. The complete nucleotide sequence was determined; it revealed a single open reading frame of 294 amino acids that included four in-frame AUG codons as potential sites of translation initiation. Comparison of the human B23 nucleotide sequence with the rat and mouse B23 sequences revealed 91% homology (hum:rat, and hum:mus) in the coding region; the predicted B23 proteins displayed 94% amino acid identity. Northern gel blot analysis revealed a single B23 mRNA species of ~1.5 kb. The level of B23 mRNA in U cells was not detectably altered by treatment with either alpha or gamma interferon. Southern gel blot analysis revealed polymorphism within the human B23 gene structure, and suggested the presence of multiple B23 genes and/or extensive splicing of B23 RNA transcripts.

Protein B23 is a major phosphoprotein component of the granular region of the nucleolus (1-4), the subcellular site of assembly of pre-ribosomal RNAs with proteins (5). Although the biochemical function of protein B23 has not yet been precisely established, the observations that protein B23 is found in the nucleolus (1-4), binds to single-stranded nucleic acids including RNA, and is associated with pre-ribosomal ribonucleoprotein

*This work was supported in part by Research Grants AI-12520 and AI-20611 from the National Institute of Allergy and Infectious Diseases, U. S. P. H. S.

[1] The nucleotide sequence reported in this paper has been submitted to GenBank, Los Alamos, New Mexico, with accession number M26697.

The abbreviations used are: IFN, interferon; B23, protein B23, numatrin; nt, nucleotide; ORF, open reading frame.

0006-291X/89 $1.50

particles (6-8) has led to the suggestion that protein B23 may play a role in ribosome biogenesis. In addition, protein B23 appears identical to numatrin (9), a nuclear matrix protein associated with growth factor-induced mitogenesis of 3T3 fibroblasts and with mitogen- and receptor-mediated proliferation of lymphocytes (9-11).

As part of our studies of the molecular mechanism of interferon action (12, 13), we have undertaken the molecular cloning of IFN-regulated human proteins. Among the cDNA clones obtained was one that turned out to be a false-positive with regard to regulation by interferon (IFN) but, upon partial sequence analysis and comparison with the GenBank data base, possessed significant sequence similarity with that of a partial cDNA clone of the human protein B23 (14) as well as that of a complete cDNA clone of rat protein B23 (7). We report here the characterization of clone 21 which appears to be a full-length molecular cDNA clone of the human nucleolar protein B23, also known as numatrin.

MATERIALS and METHODS

cDNA Synthesis and Cloning Strategy. The cDNA library in lambda Zap was prepared using poly(A^+) RNA isolated from human amnion U cells treated with IFN-α. Briefly, total RNA was isolated from monolayer cell cultures by the guanidinium thiocyanate - cesium chloride centrifugation technique (15); polyadenylated RNA was prepared by two cycles of oligo(dT) cellulose (Collaborative Research) chromatography (16). Synthesis of cDNA from poly(A^+) RNA was by the method of Gubler and Hoffman (17). Size-selected cDNA, treated with EcoRI methylase, ligated to EcoRI linkers and digested with EcoRI endonuclease, was ligated into the EcoRI site of λ-Zap expression vector (Stratagene Cloning System) according to the manufacturer's recommendations. The ligated DNA was packaged using the Gigapack Gold extract system (Stratagene) and plated onto Escherichia coli strain BB4. Initial screening was by filter hybridization (20) with a mixture of synthetic 17-mer oligonucleotides corresponding to a peptide sequence determined for protein p54, a 54-kDa polypeptide induced by IFN in human cells (18,19). Candidate positive plaques were subjected to a second round of oligonucleotide screening prior to subcloning into the Bluescript-SK plasmid (Stratagene) by the in vivo automatic excision process permitted by the λ-Zap vector. Subsequent screening was by Northern blot analysis of RNA isolated from untreated and IFN-treated cells; probes were EcoRI restriction fragment inserts isolated from Bluescript-SK recombinants and ^{32}P-labeled by random priming (21).

Northern Blot Analysis of RNA. Human U cell total RNA, isolated from untreated cells and from cells treated with either IFN-α or IFN-γ as indicated, was denatured, fractionated by electrophoresis using a 1.5% agarose-formaldehyde gel, and transferred to Hybond-N (Amersham) nylon transfer membranes (22,23). Hybridization was as described (23), using as a probe the full-length 1.3 kb EcoRI:EcoRI B23 fragment labeled by random priming.

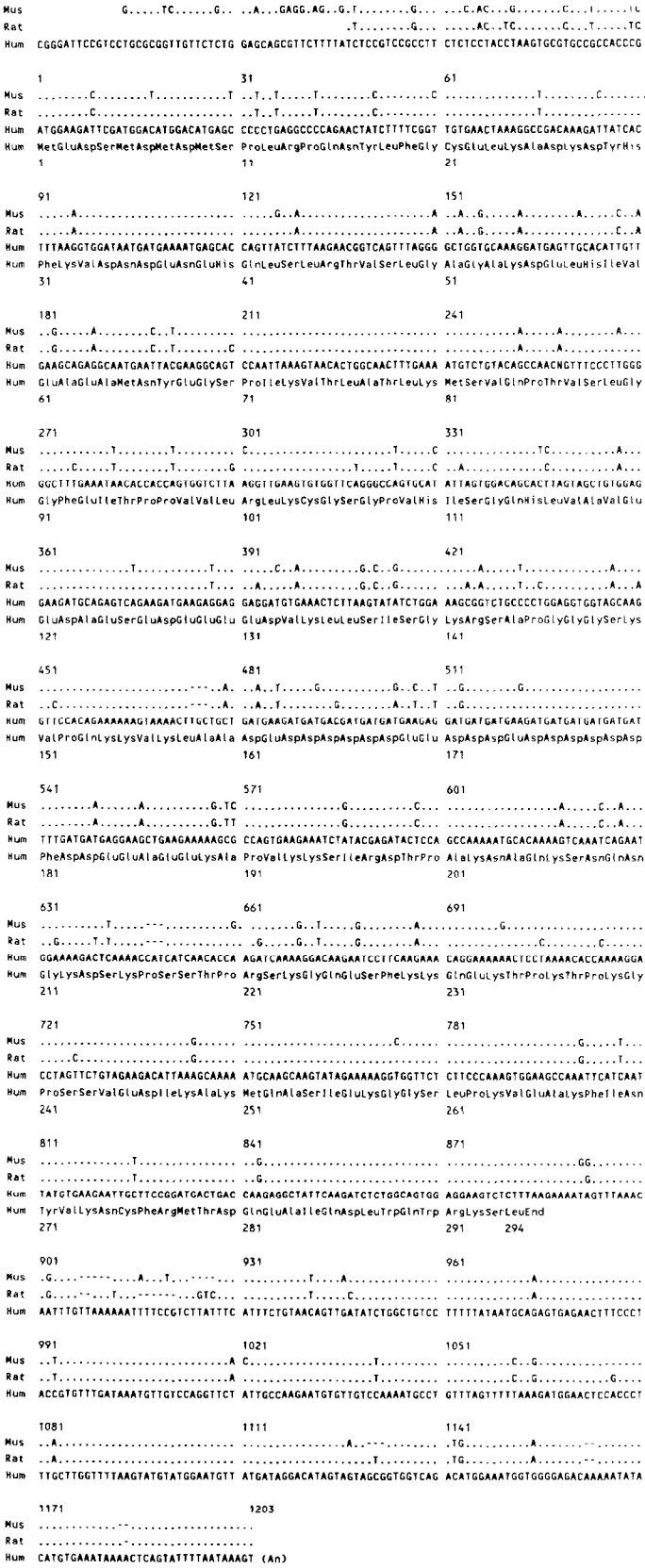
Southern Blot Analysis of Human Genome DNA. Chromosomal DNA, purified from human sperm or from Epstein-Barr virus transformed human lymphoblast lines, was digested with the restriction enzyme EcoRI or StuI (10 or 20 ug DNA, respectively), fractionated by agarose gel electrophoresis, and transferred to membranes by the method of Southern (24). The filter containing human sperm genome DNA was generously provided by M. Neitz, Univ. of Calif., Santa Barbara; the filter containing human lymphoblast DNA from CEPH families (Centre d'Etude du Polymorphisme Humain, Paris, France) was generously provided by N. Steiner and M. King, Univ. of Calif., Berkeley. The Southern blot filters were probed with the 1.3 kb EcoRI:EcoRI B23 fragment.

DNA Sequence Analysis. A family of progressive unidirectional deletions of B23 clone 21 in the Bluescript-SK+ vector was generated using exonuclease III (25). Sequencing was done with the dideoxy chain termination method (26) using modified T7 DNA polymerase (Sequenase II, United States Biochemical Corp.), ^{35}S-labeled deoxy-ATP, and the T7 and T3 primers according to the Sequenase protocol provided by the manufacturer. In some cases, the reaction mixtures included dITP in place of dGTP to minimize band compressions. Computer assisted sequence analysis was performed with a Digital Equipment Corp. VAXSTATION II computer system and the sequence analysis software package of the Univ. of Wis. Genetics Group (33); comparisons of sequences were made against Genbank release 59.0.

RESULTS and DISCUSSION

Nucleotide Sequence of Human B23 (Numatrin) cDNA. The nucleotide sequence of cDNA clone 21 isolated from a lambda-Zap cDNA library constructed using poly(A^+) RNA from IFN-α treated human U cells is shown in Figure 1. The clone was identified by filter hybridization with a mixture of synthetic oligonucleotides complementary to the likely mRNA codons encoding a short peptide sequence derived for IFN-regulated protein p54 (18,19). Clone 21 turned out to be a false-positive, both with regard to p54 and with regard to regulation by IFN (Fig. 3).

Figure 1. Nucleotide sequence of human cDNA clone 21 and deduced amino acid sequence of human protein B23 encoded by clone 21. The numbers above each line refer to nucleotide position and numbers beneath each line to amino acid position. The A of the 5'-proposed AUG initiator methionine codon of the long open reading frame of clone 21 is designated as nucleotide +1; nucleotides -90 to -1 designate the predicted 5'-untranslated region. About 40 A residues were present at the 3'-end of clone 21 (sequence not shown) and are designated by $(A)_n$. For purposes of comparison with the human B23 (hum) cDNA clone 21 sequence (accession number M26697), the complete cDNA sequences determined for rat B23 (rat) by Chang et al (7; accession number J03969) and mouse B23 (Mus) by Schmidt-Zachmann and Franke (30) are also included. Dots (.) designate identity of nucleotides in the compared rat or mouse sequence with those of human B23. Dashes (-) designate gaps introduced to maximize similarity.



Clone 21 was, however, identified as human protein B23, numatrin (9), based on sequence similarity to the DNA sequence reported for rat protein B23 (7) and a partial sequence reported for a version of the human protein B23 (14).

Human cDNA clone 21 is 1293 nucleotides of heteropolymeric sequence followed by a 3′ poly A tail (Fig. 1). Clone 21 possesses a single, long open reading frame of 882 nt, capable of encoding a 294 amino acid protein. The ORF is preceded by 90 nt which presumably constitute a part of the 5′-untranslated leader. The 5′-proximal AUG codon of the ORF is designated nt 1-3; this potential translation initiation codon is closely followed by three additional in-frame AUG codons at nt positions 13-15, 19-21, and 25-27. It is unclear which of these AUG codons represents the major translation initiation site of B23. The AUG codons at nt 1-3 and nt 13-15 are not flanked by a favorable "-3/+4" context for translation initiation, whereas the AUG codons at nt 19-21 and nt 25-27 possess purines at the "-3/+4" positions and hence would be considered more favorable (27). A number of different animal virus genes encode mRNA species in which different in-frame AUG initiation sites are utilized to generate related protein products that differ in their amino-terminal sequences (28). Different forms of B23 protein have been described (6,14,29); conceivably the heterogeneity of protein B23 is due in part to the utilization of alternative, in-frame translation initiation sites.

The 3′-untranslated region of human clone 21 B23 mRNA is relatively long, consisting of 321 nucleotides. Two copies of the polyadenylation signal AAUAAA, located at nt 1178-1183 and nt 1196-1201, are closely positioned to each other and to the 3′-terminus preceding the poly A tail (Fig. 1).

```
     1                                                  51
Hum  MEDSMDMDMS PLRPQNYLFG CELKADKDYH FKVDNDENEH QLSLRTVSLG AGAKDELHIV EAEAMNYEGS PIKVTLATLK MSVQPTVSLG GFEITPPVVL
Rat  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Mus  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........

     101                                                151
Hum  RLKCGSGPVH ISGQHLVAVE EDAESEDEEE EDVKLLSISG KRSAPGGGSK VPQKKVKLAA DEDDDDDDEE DDDEDDDDDD FDDEEAEEKA PVKKSIRDTP
Rat  .......... .......... ........D. ......GM.. ........N. ........DE .D.E..E.D. -.......... ..E..T...V .....V....
Mus  .......... .......... ........D. ......GM.. ........N. ........DE .DE...E.D. -.......... ..E..T...V .....V....

     201                                                251                                                294
Hum  AKNAQKSNQN GKDSKPSSTP RSKGQESFKK QEKTPKTPKG PSSVEDIKAK MQASIEKGGS LPKVEAKFIN YVKNCFRMTD QEAIQDLWQW RKSL
Rat  .......... ...L..-... .......... .......... .......... .......... .......... .......... .......... ....
Mus  .......... ...L..-... .......... .......... .......... .......... .......... .......... .......... ....
HB1              P..S..SS S.......... .......... .......... .......... .......... .......... .......... ....
HB2              P........ .......... .......... .......... .......... .......... .......... .......... ....
```

Figure 2. Comparison of the amino acid sequences deduced from molecular clone cDNA sequences for human protein B23 (Hum), rat protein B23 (Rat), mouse protein B23 (Mus) and human B23 proteins HB1 and HB2 (HB1, HB2). Sequences were aligned by computer analysis and visual inspection. Dots (.) designate identity of residues in the compared sequence with those of human B23. Dashes (-) designate gaps introduced to maximize similarity. Hum, human protein B23 sequence deduced from clone 21 (Figure 1); Rat, rat protein B23 sequence deduced from clone λJH1 (7); Mus, mouse protein B23 deduced from clone λFML185.19 (30); HB1 and HB2, human protein B23 carboxyl terminal region deduced from partial cDNA clones hpB1 and hpB2 (14).

Relatedness of the Human, Rat and Mouse B23 (Numatrin) Genes and Encoded Polypeptides. The human, rat and mouse B23 genes and encoded proteins are highly conserved (Fig. 2). Comparison of the B23 protein sequence predicted from the human clone 21 with the rat (7) and mouse (30) B23 protein sequences deduced from cDNA clones revealed 94% amino acid identity for both hum:rat and for hum:mus. The carboxyl-terminal sequence of human protein B23 deduced from U cell clone 21 differs from the sequences of human B23 predicted from two partial placenta cDNA clones, HB1 and HB2 (14), by five and one amino acids, respectively (Fig. 2). As reported previously for the rat (7) and mouse (30) B23 proteins, the human B23 protein (Figs. 1 and 2) likewise possesses extensive sequence similarity with the Xenopus protein N038 (31).

Among the more intriguing structural domains displayed by the human B23 protein encoded by clone 21 is a negatively charged domain corresponding to amino acid residues 161 to 188. Within

this domain, 26 of 28 codons are GAN and specify Asp or Glu; furthermore, 17 of the 18 Asp codons are GAU, and 6 of the 8 Glu codons are GAG (Fig. 1).

Comparison of the human clone 21 B23 nucleotide sequence with the cDNA nucleotide sequences reported for rat (7) and mouse (30) B23 clones revealed 91% nucleotide homology for both hum:rat and for hum:mus in the coding region (Fig. 1). The 3'-untranslated sequences possessed a comparable homology, 90% for hum:rat and 89% for hum:mus. However, the 5'-untranslated sequences displayed a somewhat lower homology, 74% for hum:rat and 71% for hum:mus. Following the completion of the clone 21 cDNA sequence corresponding to human amnion U cell B23, we became aware of the complete cDNA sequence for human placenta B23 recently reported by Chan et al. (32); the U cell (Fig. 1) and placenta (32) human B23 cDNA sequences are equivalent.

Northern Gel Blot Analysis of Human B23 (Numatrin) mRNA in Untreated and IFN-treated Cell. Northern blot analysis of total RNA isolated from human amnion U cells revealed a single mRNA transcript with an electrophoretic mobility slightly faster than 18S rRNA (Fig. 3A,B). The size of the U cell B23 transcript was estimated to be about 1.5 kb. The level of the B23 transcript was comparable in U cells not treated with IFN and in cells treated with alpha IFN, gamma IFN, or a combination of IFNs (Fig. 3A,B).

Genomic Southern Analysis of the Human B23 (Numatrin) Gene. To study the molecular organization of the human B23 gene, human genomic DNA digested with either EcoRI or StuI was analyzed, by the method of Southern (24), for hybridization to the clone 21 cDNA probe of human U cell B23. No EcoRI or StuI restriction sites are present within the human B23 clone 21 cDNA (Fig. 1). Human sperm DNA obtained from two individuals, when digested with EcoRI, yielded multiple fragments that hybridized under stringent

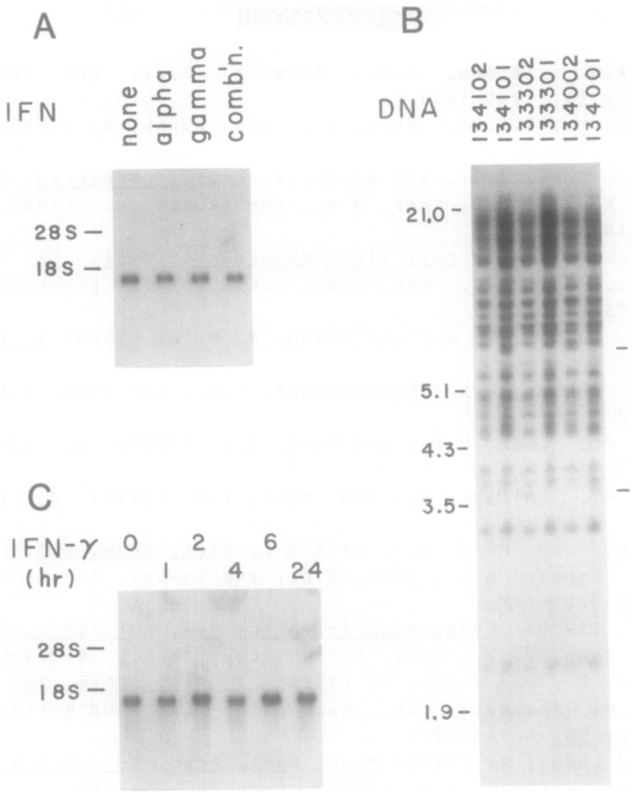


Figure 3. (Fig. 3A,B) Northern blot analysis of B23 mRNA expression in untreated and interferon-treated human amnion U cells. U cell monolayer cultures were untreated or treated with IFN-α or IFN-γas indicated. RNA was purified from cells, denatured and electrophoresed on an agarose gel, blotted onto a nylon membrane, and probed with ^{32}P-labeled cDNA to human B23 message as detailed under "Materials and Methods." (A) Untreated (none) or treated with alpha IFN (α), gamma IFN (γ), or alpha + gamma IFN (comb'n) for 24 h; (B) Treated with IFN-γ for 0, 1, 2, 4, 6, or 24 h. (Fig. 3C) Southern blot analysis of human genomic DNA. Human genomic DNA isolated from Epstein-Barr virus transformed lymphoblasts of six CEPH individuals was digested with StuI, blotted and probed as described. Sizes of DNA standards, in kilobases, are given on the left, and positions of polymorphism on the right. The blots shown in (A), (B) and (C) were hybridized with a ^{32}P-labeled full-length 1.3 kb EcoRI:EcoRI insert of human B23 clone 21.

wash conditions to human B23 clone 21 cDNA (data not shown). Similar results were obtained when human lymphoblast genomic DNA from six CEPH individuals was analyzed (Fig. 3C). Multiple genomic StuI fragments hybridized with the clone 21 probe. Furthermore, StuI site polymorphism was observed with human genomic DNA hybridizing to the clone 21 probe (Fig. 3C). These results suggest that the human B23 gene is interrupted by several intervening sequences, and/or that the human genome includes multiple genes related to the human B23 cDNA clone 21.

REFERENCES

1. Prestayko, A.W., Klomp, G.R., Schmoll, D.J., and Busch, H. (1974) Biochemistry 13, 1945-1951.
2. Kang, Y.J., Olson, M.O.J., Jones, C., and Busch, H. (1975) Cancer Res. 35, 1470-1475.
3. Spector, D.L., Ochs, R.L., and Busch, H. (1984) Chromosoma 90, 139-148.
4. Chan, P.-K., Aldrich, M., Cook, R.G., and Busch, H. (1986) J. Biol. Chem. 261, 1868-1872.
5. Hadjiolov, A. A. (1985) Cell Biol. Monographs, vol. 12 Springer, Berlin.
6. Yung, B.Y.M., Busch, H., and Chan, P.K. (1986) Biochim. Biophys. Acta 826, 167-173.
7. Chang, J.-H., Dumbar, T. S., and Olson, M. O. J. (1988) J. Biol. Chem. 263, 12824-12827.
8. Borer, R.A., Hehner, C.F., Eppenberger, H.M., and Nigg, E.A. (1989) Cell 56, 379-390.
9. Feuerstein, N., Chan, P.K., and Mond, J.J. (1988) J. Biol. Chem. 263, 10608-10612.
10. Feuerstein, N., Spiegel, S., and Mond, J.J. (1988) J. Cell Biol. 107, 1629-1642.
11. Feuerstein, N., and Mond, J.J. (1987) J. Biol. Chem. 262,11389-11397.
12. Pestka, S., Langer, J.A., Zoon, K.C., and Samuel, C.E. (1987) Annu. Rev. Biochem. 56, 727-777.
13. Samuel, C. E. (1988) Prog. Nucleic Acids Res. Mol. Biol. 35, 27-72.
14. Chan, P.-K., Chan, W.-Y., Yung, B.Y.M., Cook, R.G., Aldrich, M.B., Ku, D., Goldknop, I. L., and Busch, H. (1986) J. Biol. Chem. 261, 14335-14341.
15. Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J., and Rutter, W.J. (1979) Biochemistry 18, 5294-5299.
16. Aviv, H., and Leder, P. (1972) Proc. Natl. Acad. Sci. U.S.A. 69, 1408-1412.
17. Gubler, U., and Hoffman, B.J. (1983) Gene 25, 263-269.
18. Ulker, N., and Samuel, C.E. (1985) J. Biol. Chem. 260, 4324-4330.
19. Ulker, N., Zhang, X.,and Samuel, C.E. (1987) J. Biol.Chem. 262, 16798-16803.
20. Karathanasis, S.K. Norum. R.A., Zannis, V.I., and Breslow, J.L. (1983) Nature 301, 718-720.
21. Feinberg, A.P., and Vogelstein, B. (1983) Anal. Biochem. 132, 6-13.
22. Thomas, P.S. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 5201-5205.
23. Atwater, J.A., Munemitsu, S.M., and Samuel, C.E. (1987) Virol. 159, 350-357.
24. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
25. Henikoff, S. (1987) Methods Enzymol. 155, 156-165.
26. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463-5467.
27. Kozak, M. (1984) Nucleic Acids Res. 12, 857-871.
28. Samuel, C.E. (1989) Prog. Nucleic Acids Res. Mol. Biol. 37, 127-153.
29. Chan, P.-K., Aldrich, M.B.,and Busch, H. (1985) Exp. Cell Res. 161, 101-110.
30. Schmidt-Zachmann, M.S., and Franke, W.W. (1988) Chromosoma 96, 417-426.
31. Schmidt-Zachmann, M.S., Hugle-Dorr, B., and Franke, W.W. (1987) EMBO J. 6, .
32. Chan, W.-Y., Liu, Q.-R, Borjigin, J., Busch, H., Rennert, O.M., Tease, L.A., and Chan, P.-K. (1989) Biochemistry 28, 1033-1039.
33. Devereus, J., Haeberli, P., and Smithies, O. (1984) Nucleic Acids Res. 12, 387-395.