Structure of the mouse *klotho* gene and its two transcripts encoding membrane and secreted protein

Takako Shiraki-Iida[a,b], Hiroki Aizawa[a,c], Yutaka Matsumura[a,c], Susumu Sekine[b], Akihiro Iida[b], Hideharu Anazawa[b], Ryozo Nagai[c,d], Makoto Kuro-o[a,d,*], Yo-ichi Nabeshima[a,d,e]

[a] *Division of Molecular Genetics, National Institute of Neuroscience, NCNP, 4-1-1 Ogawahigashi, Kodaira, Tokyo 187, Japan*
[b] *Tokyo Research Laboratories, Kyowa Hakko Kogyo Co., Ltd., 3-6-6 Asahimachi, Machidashi, Tokyo 194, Japan*
[c] *The 2nd Department of Internal Medicine, Faculty of Medicine, University of Gunma, 3-39-22 Syowa, Maebashi, Gunma 371, Japan*
[d] *Core Research for Evolutional Science and Technology (CREST), JRDC, Tokyo, Japan*
[e] *Institute for Molecular and Cellular Biology, Osaka University, 1-3 Yamada-oka, Suita, Osaka 565, Japan*

Received 5 January 1998; revised version received 27 January 1998

Abstract We previously established a novel mouse model for human aging and identified the genetic foundation responsible for it. A defect in expression of a novel gene, termed *klotho* (*kl*), leads to a syndrome resembling human aging in mice. The *kl* gene encodes a single-pass membrane protein whose extracellular domain carries homology to β-glucosidases. In this report, we present the entire mouse *kl* gene organization. The mouse *kl* gene spans about 50 kilobases and consists of five exons. The promoter region lacks a TATA-box and contains four potential binding sites for SP1. We further show that two *kl* gene transcripts encoding membrane or secreted protein are generated through alternative transcriptional termination. These findings provide fundamental information for further study of the *kl* gene which may regulate aging in vivo.

Key words: Mouse *klotho* gene; Alternative transcriptional termination; Competitive polymerase chain reaction

1. Introduction

Recently, we have identified a novel gene, termed *klotho* (*kl*), that is involved in the development of a syndrome resembling human aging in mice [1]. A defect in *kl* gene expression in mice leads to multiple disorders including arteriosclerosis, osteoporosis, pulmonary emphysema and skin atrophy together with short life-span and infertility. The *kl* gene encodes a novel single-pass membrane protein of 1014 amino acids, whose extracellular domain consists of two-fold internal repeats with homology to β-glucosidases [2,3]. The *kl* gene transcript was not detected in many organs, including the stomach, lung, submandibular gland, skin and bone, where severe changes occurred in the mutants. To explain these apparently cell non-autonomous phenomena, it must be assumed that a secreted circulating factor mediates the pleiotropic functions of the Klotho (KL) protein, at least in part [1].

In this article, we report the structure of the mouse *kl* gene and its promoter region and the identification of a novel variant of *kl* gene transcript that encodes a putative secreted form of KL protein. The discovery of a secreted form of *kl* gene product may provide important information for understanding the molecular function of the *kl* gene and the molecular mechanisms of human aging and aging related disorders.

*Corresponding author (a). Fax: (81) (423) 46-3864.
E-mail: kuroo@ncnaxp.ncnp.go.jp

The nucleotide sequence data reported in this paper will appear in the DDBJ, EMBL and GenBank nucleotide sequence databases with the accession numbers AB010088, AB010089, AB010090 and AB010091.

2. Materials and methods

2.1. Isolation of the mouse kl gene

A bacterial artificial chromosome (BAC) library (Research Genetics Inc., Huntsville, AL) was screened by PCR using the primer pair 5′-AGGCTCATCAGAGGAACTGC-3′ and 5′-AAGACAGAAGCTGCCTCAGG-3′ specific for mouse *kl* cDNA. The positive BAC clones were partially digested with *Sau*3AI and subcloned into the cosmid vector pWE15 (Clontech) at the *Bam*HI site to construct a cosmid contig of the mouse *kl* locus. Intron size was determined by sequencing or estimated by PCR using primers specific for the flanking exons. Sequencing was performed using the cycle sequencing method with SequiTherm EXCEL II DNA Polymerase (SequiTherm EXCEL Long-Read DNA Sequencing Kits-LC, Epicentre Technologies) in a LI-COR Model 4000L Automated DNA Sequencer according to the manufacturer's instructions. Sequence data were analyzed using GeneWorks software (Oxford Molecular Group, Inc.).

2.2. Determination of transcription start sites

An S1 nuclease protection assay was performed using an S1 mapping kit (Ambion) according to the supplier's protocol. A 395 bp genomic DNA fragment containing a putative transcription initiation site was synthesized by PCR using the primer pair 5′-CCTGCTGTGCTCTCTCTGGG-3′ and 5′-AAACCGTCGGGGAAGGTGT-3′. The fragment was 5′ end-labeled using the Megalabel Kit (TaKaRa). The labeled DNA was strand separated and the antisense strand was hybridized with 6 μg mouse kidney poly(A)$^+$ RNA at 42°C for 16 h, then digested with 500 U S1 nuclease at 37°C for 30 min. The protected fragments were analyzed by electrophoresis on a 5% acrylamide/8 M urea gel.

2.3. RT-PCR

Poly(A)$^+$ RNA (500 ng) from various organs was reverse transcribed with random hexamer (TaKaRa), and 5% of the reaction mixture was amplified by LA-Taq DNA polymerase (TaKaRa) using a primer pair specific for the mouse secreted isoform (*INS-2* (Fig. 3): 5′-GGGTTTTGTCAAAGGACTTAC-3′ and mEX6: 5′-AGGCTCATCAGAGGAACTGC-3′ in exon 2 of the *kl* gene). The condition for amplification was 30 cycles of 94°C for 20 s, 60°C for 30 s, 72°C for 60 s.

2.4. 3′ RACE-PCR and 5′ RACE-PCR

A clone containing the 3′ untranslated region of the mouse secreted isoform was generated by the rapid amplification of cDNA end method using the Mouse Kidney Marathon-Ready cDNA kit (Clontech) [4,5]. For the primary PCR reaction, the adaptor primer-1 (AP1) provided in the kit and the forward primer *INS-5U* (Fig. 3) (5′-GACAAAACCCAGTGTCGGCCTC-3′) were used. For the secondary PCR, the adaptor primer-2 (AP-2 nested to AP-1) provided in the kit and the forward primer *INS-6U* (Fig. 3) (5′-CACTAAGCTCTGGCCAAGGCAC-3′) were used. Using the same kit, the 5′ end of a

PII S0014-5793(98)00127-6

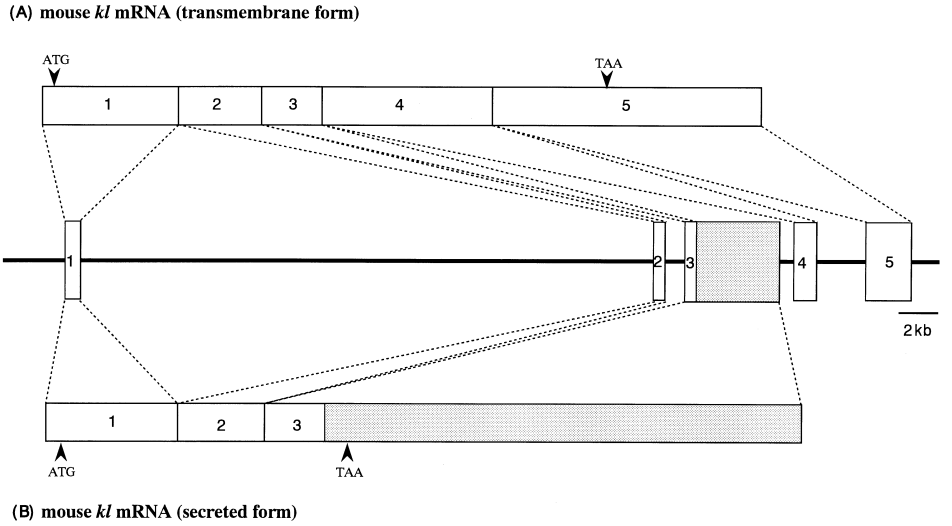

Fig. 1. Structure of the mouse *kl* gene and its two gene transcripts. Open boxes represent exons and the thick line represents introns. A: The transcript encoding the membrane protein is 5.2 kbp mRNA and contains five exons. B: The transcript encoding the secreted protein is 5.8 kbp mRNA and contains three exons. The gray box represents the continuation of exon 3 in the secreted isoform. The translation start codon and the termination codon are shown above the transmembrane form and under the secreted form transcripts, respectively.

mouse secreted isoform cDNA was amplified. The first PCR was performed using AP-1 and *INS-1* (Fig. 3) (5′-TTAGTGAGGAAG-CAAGAGGCC-3′). The secondary PCR was performed using AP-2 and *INS-2*. The secondary PCR product was directly cloned into the pCR2.1 vector (Invitrogen) and sequenced.

2.5. Competitive PCR

Specific competitors of the membrane or secreted isoform were obtained using PCR MIMIC Construction Kit (Clontech). Poly(A)$^+$ RNA (500 ng) from various tissues was reverse transcribed with random hexamer (TaKaRa), and 5% (v/v) of the reaction mixture and competitor was amplified by EX-Taq DNA polymerase (TaKaRa). The conditions for PCR were 25 cycles of 94°C for 20 s, 60°C for 30 s and 72°C for 1 min. PCR products were analyzed by 1.5% agarose gel electrophoresis. In the competitive PCR of the secreted form transcripts, the 549 bp and 321 bp PCR products represent the secreted isoform and its competitor, when the same primer set for RT-PCR was used. In the PCR of the membrane isoform transcripts, the 597 bp and 323 bp PCR products represent the membrane form and its competitor, when mEX8R (5′-CCTCTTACTGTGATGCACA-TCC-3′ in exon 4 of the *kl* gene) instead of the *INS-2* primer was used. A photograph of the illuminated gel (Polaroid 667 film) was scanned using PDI-Quantity One (Toyobo). The amount of mouse *kl* mRNA was calculated following the procedure of the PCR MIMIC Construction Kit (Clontech) [6–8].

3. Results

A total of five independent BAC clones were isolated by PCR screening of a mouse BAC genomic library [9]. These clones harbored approximately 100 kbp inserts estimated by pulse-field gel electrophoresis (data not shown). Sequence analysis of the BAC clones revealed that the mouse *kl* gene spans about 50 kbp and consists of five exons (Fig. 1). Exon-intron junctions and the distance between exons are shown in Table 1. S1-nuclease mapping analysis revealed the existence

Table 1
Exon/intron characteristics of the mouse *kl* gene

Exon	3′ Splice site			5′ Splice site	Exon size	Intron size
		838				
1		CTA CTT TTG		**gt**tgagttcga	838 bp	>33 kbp
		L L L				
		839	1349			
2	tcctctgt**ag**	GCT CAT GCC	... ACC TTA AAA G	**gt**aggctttc	511 bp	1372 bp
		A H A	**T L K**			
		1350	1618			
3	ccctctgc**ag**	CA ATC AHA	... TAC GTT CAA	**gt**aagtcctt	269 bp	5604 bp
		A I R	**Y V Q**			
		1619	2720			
4	ttctttgt**ag**	GTG GAC ACT	... GCT CTG AAA G	**gt**gagacccc	1102 bp	>1.8 kbp
		V D T	**A L K**			
		2721				
5	ttttctct**ag**	CC TAC GTG			2299 bp	
		A Y V				

Exon and intron sequences are indicated by uppercase and lowercase letters, respectively. Amino acids are indicated by bold letters. Nucleotide position is assigned starting from the major transcription start site. The gt-ag (5′→3′) rule of the splice site is conserved. Intron size was determined by direct sequencing (intron 2 and 3) or estimated by PCR.

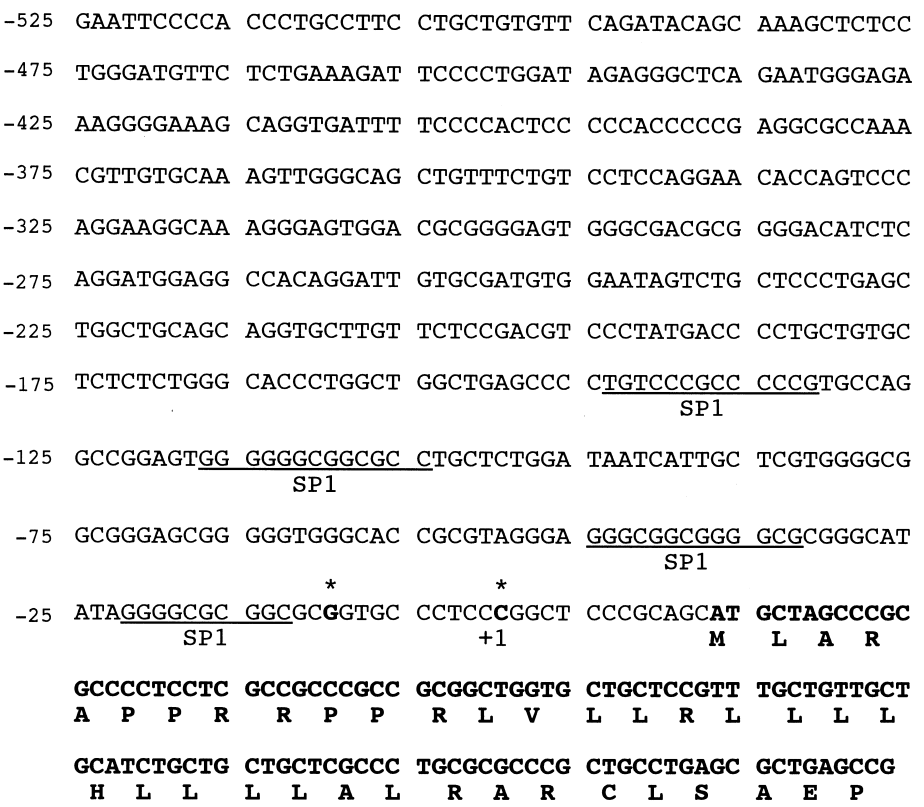

Fig. 2. Promoter region of the mouse *kl* gene. The nucleotide sequence of the exon is indicated by bold letters. The 5′ untranslated sequence is indicated by uppercase letters. The two transcription start sites are indicated by asterisks, and the major transcription start site is designated +1. Four putative SP1 binding sites are underlined.

of two transcription start sites. The major start site is mapped to 13 bp upstream of the translation start (ATG) and the minor start site to 23 bp upstream. In the 5′ flanking region up to 540 bp upstream from the major transcription start site, there are neither TATA-like consensus sequences nor CAAT *cis*-acting elements; there are, however, four putative SP1 binding sites (Fig. 2) [10]. In the human *kl* gene, we have found that alternative splicing at an internal splice donor site in exon 3 results in two splice variants, which encode either a membrane or a secreted form of KL protein [11]. In order to determine whether there are similar splice variants in mice, we examined mouse kidney cDNA using primers specific for mouse genomic sequences in the vicinity of exon 2, exon 3 or exon 4. As a result, a second type transcript encoding the

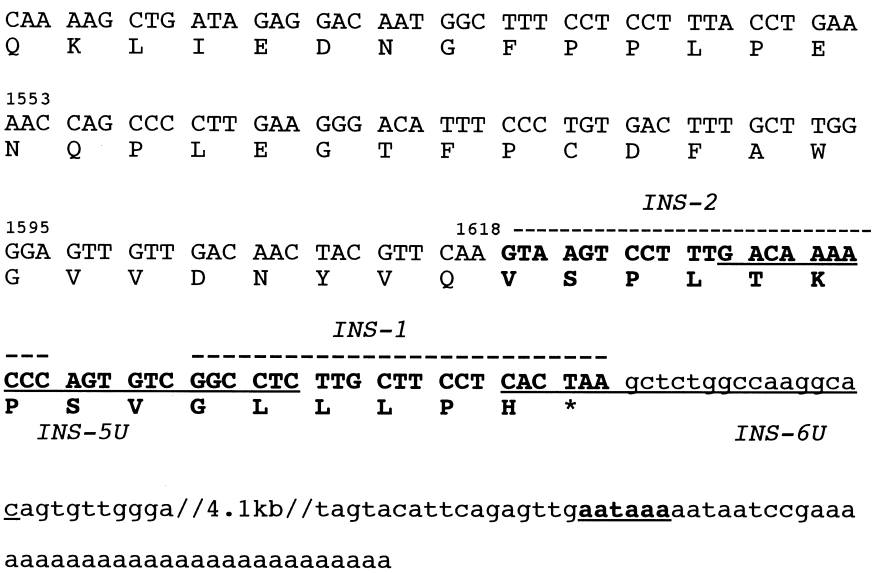

Fig. 3. Partial cDNA sequence of the secreted isoform of the mouse *kl* gene. The predicted amino acid sequence is shown beneath the cDNA sequence in single letter code. The termination codon is indicated by an asterisk. Translated sequences specific for the secreted isoform are indicated by bold letters. The primer sites used for the 3′ RACE and 5′ RACE analysis are indicated by dotted lines and underlines, respectively. The putative polyadenylation signal is indicated by underlined bold letters.

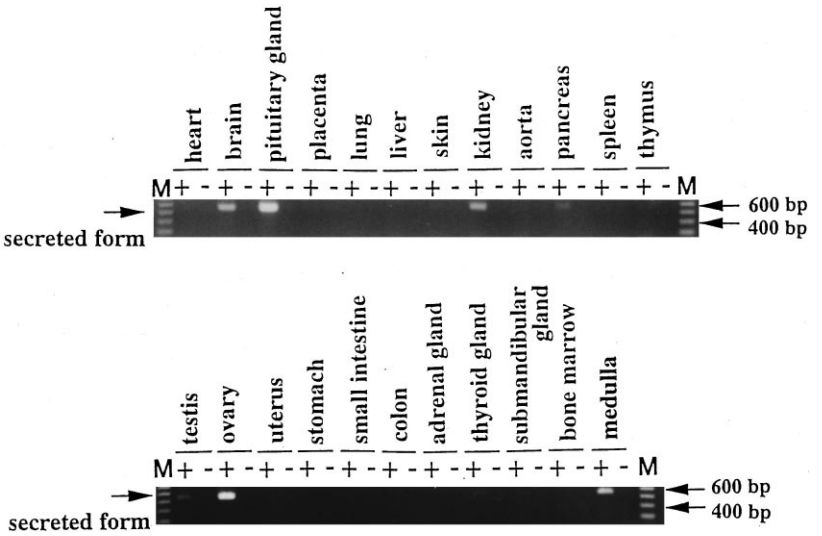

Fig. 4. Expression of the secreted isoform of the mouse *kl* gene. Expression of the secreted isoform of the mouse *kl* gene was detected by RT-PCR in various tissues. The (+) lanes show the samples reacting with AMV reverse transcriptase, and (−) lanes are negative controls (no reverse transcriptase). The size of the expected fragment (549 bp) is indicated by an arrow. Expression of GAPDH was examined in all tissues (data not shown). M: 100 bp ladder marker.

secreted form of KL protein was identified (Fig. 1B). Precursor RNA of the second type transcript initiates from the same start sites as the membrane form transcript and terminates about 4.2 kbp downstream of the 3′ end of exon 3 of the membrane form (Fig. 1). The transcriptional termination site is located near upstream of exon 4 of the membrane form and thus the splicing acceptor sequence of intron 3 is excluded from the precursor transcript of the secreted form, making the long stretch of sequence corresponding to intron 3 of the membrane form to be unspliced from the mature secreted form mRNA. As a result, the secreted form of the *kl* gene product is 5.8 kbp in length and composed of exons 1, 2 and long exon 3. A polyadenylation signal (AATAAA) is found near its end (Fig. 3) and an in-frame translation stop codon (TAA) appeared 49 bp downstream of the 3′ end of exon 3 of the membrane form transcript (Figs. 1 and 3). The second type of *kl* transcript encodes a putative secreted protein of 550 amino acids (corresponding to mKL1 [1]), lacking the transmembrane domain.

Expression of secreted form mRNA is not detected by Northern blot analysis, but it is detected by RT-PCR in several tissues. As shown in Fig. 4, the secreted form is expressed in the pituitary gland, ovary, whole brain, brain medulla, kidney, testis, pancreas and thyroid gland. Expression was not detected even by RT-PCR in other tissues. The tissue distribution of the secreted form is similar to that of the membrane form [1]. Competitive PCR revealed that expression of the membrane form is about 10 times greater than the secreted form in the pituitary gland (Fig. 5). In other tissues, the membrane form always predominates the secreted form by more than 10 times (data not shown). The ratio of secreted form transcript to membrane form transcript in various tissues does not significantly change with age (data not shown).

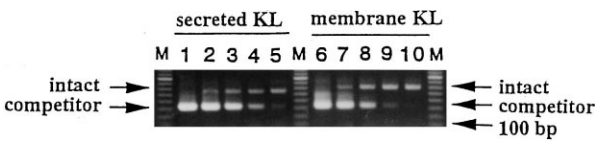

Fig. 5. Quantitation of the expression of the two isoforms by RT-competitive PCR in pituitary gland. The amount of competitor is 10 amol/μl in lanes 1 and 6, 1.0 amol/μl in lanes 2 and 7, 10^{-1} amol/μl in lanes 3 and 8, 10^{-2} amol/μl in lanes 4 and 9, 10^{-3} amol/μl in lanes 5 and 10. Lanes 1–5 contain the secreted isoform, and lanes 6–10 contain the membrane isoform. Three independent experiments were performed. The secreted isoform expressed 1.9×10^5 copies/μg mRNA, and the membrane isoform expressed 1.9×10^6 copies/μg mRNA. M: 100 bp ladder marker.

4. Discussion

In this study, we have determined the structure of the mouse *kl* gene and its promoter. The promoter region contains several potential Sp1 binding sites and shows characteristics of a TATA-less promoter. We further identified two transcripts encoding either a membrane or a secreted protein. These results provide fundamental information for further investigation of the *kl* gene. The *kl* mutation does not disrupt the coding structure of the *kl* gene but severely reduces its expression [1]. The structural analysis of the *kl* gene revealed that the insertional mutation in *kl/kl* mice is located not in the *kl* gene but in the 5′ flanking region about 6 kbp upstream of the transcription start site. This may be the reason why the *kl* gene is slightly transcribed in the mutants. Since the mutation is accompanied by a deletion of about 8 kbp, it could be that *cis*-acting element(s) with enhancer activity exist within the deletion.

The transcript of the membrane isoform contains five exons, showing the same organization in both mice and humans [11]. On the other hand, the transcript of the mouse secreted isoform contains three exons, whereas the transcript of the human secreted isoform contains five exons. In addition, the expression level of the two isoforms is very different between mice and humans: the membrane isoform predominates over the secreted isoform in mice, whereas it is vice versa in humans. To assume the circulating factor mediating the functions of Klotho protein, we should assume at least

three possible hypotheses: (1) Klotho protein converts an inactive precursor to an active humoral factor, (2) Klotho protein itself is secreted into extracellular spaces and function as a humoral factor, or (3) Klotho protein works as a receptor and mediates the signal required for the production of a novel humoral substance. Further investigation is necessary to determine how KL protein is involved in the suppression of a syndrome resembling aging.

Acknowledgements: We thank Dr. Seiga Itoh for his support.

References

[1] Kuro-o, M., Matsumura, Y., Aizawa, H., Kawaguchi, H., Suga, T., Utsugi, T., Ohyama, Y., Kurabayashi, M., Kaname, T., Kume, E., Iwasaki, H., Iida, A., Shiraki-Iida, T., Nishikawa, S., Nagai, R. and Nabeshima, Y. (1997) Nature 390, 45–51.
[2] Mantei, N., Villa, M., Enzler, T., Wacker, H., Boll, W., James, P., Hunziker, W. and Semenza, G. (1988) EMBO J. 7, 2705–2713.
[3] Grabnitz, F., Seiss, M., Rucknagel, K.P. and Staudenbauer, W.L. (1991) Eur. J. Biochem. 200, 301–309.
[4] Frohman, M.A., Duch, M.K. and Martin, G.R. (1988) Proc. Natl. Acad. Sci. USA 85, 8998–9002.
[5] Edwards, J.B.D.M., Delort, J. and Mallet, J. (1991) Nucleic Acids Res. 19, 5227–5232.
[6] Horikoshi, T., Danenberg, K.D., Stadlbauer, T.H.W., Volkenandt, M., Shea, L.C.C., Aigner, K., Gustavsson, B., Leichman, L., Froesing, R., Ray, M., Gibson, N.W., Spears, C.P. and Danenberg, P.V. (1992) Cancer Res. 52, 108–116.
[7] Van den Heuvel, J.P., Tyson, F.L. and Bell, D.A. (1993) BioTechniques 14, 395–398.
[8] Siebert, P.D. and Kellogg, D.E. (1995) in: The Practical Approach Series, PCR 2, pp. 135–148, Oxford University Press, Oxford.
[9] Shizuya, H., Birren, B., Kim, U.J., Mancino, V., Slepak, T., Tachiiri, Y. and Simon, M. (1992) Proc. Natl. Acad. Sci. USA 89, 8794–8797.
[10] Briggs, M.R., Kadonaga, J.T., Bell, S.P. and Tjian, R. (1986) Science 234, 47–52.
[11] Matsumura, Y., Aizawa, H., Shiraki-Iida, T., Kuro-o, M. and Nabeshima, Y. (1998) Biochem. Biophys. Res. Commun., in press.