Analysis of the promoter of the MUC1 gene overexpressed in breast cancer

Joseph Z. Zaretsky[a,*], Ronit Sarid[b], Yael Aylon[a], Leonid A. Mittelman[c], Daniel H. Wreschner[a], Iafa Keydar[a]

[a] *Department of Cell Research and Immunology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Ramat Aviv 69978, Israel*
[b] *Department of Human Microbiology, Sackler School of Medicine, Tel Aviv University, Ramat Aviv, Israel*
[c] *Interdepartmental Core Facility, Sackler School of Medicine, Tel Aviv University, Ramat Aviv, Israel*

Received 3 October 1999

Abstract The MUC1 gene encodes a mucin glycoprotein and is overexpressed in breast cancer. Knowledge of the mechanisms leading to MUC1 overexpression may help in the development of molecular approaches for breast cancer therapy. In order to study the regulation of the MUC1 gene transcription, we analyzed functional activities of various deletion mutants of the MUC1 promoter. We established that transcriptional *cis*-elements present in the *Sac*I/*Xmn*I fragment of the promoter are competent and sufficient for expression of, at least, tandem repeats containing isoform(s) of the MUC1 protein. CAT transfection analysis showed that both the 3′ and 5′ regions of the *Sac*I/*Xmn*I fragment possess transcription activities. Promoter activities associated with the *Sac*I/*Xmn*I fragment were confirmed by a RNase protection assay, which demonstrated multiple transcription start sites (TSSs) in the MUC1 gene transcribed in epithelial T47D cells. We show that treatment of the T47D cells with TGFβ1 leads to activation of additional TSSs in the MUC1 gene. The roles of the structural and functional properties of the MUC1 promoter in MUC1 gene transcription are discussed.

Key words: MUC1 gene; Promoter; CAT assay; Transcription start site

1. Introduction

MUC1 is a glycoprotein which is expressed at a basal level in most epithelial cells [1]. In breast cancer cells, its expression is dramatically increased [1–3]. To date, several isoforms of the MUC1 gene product have been identified and characterized [4–8]. The MUC1/REP isoform contains a 20 amino acid tandem repeat array and participates in cell to cell interactions [4,9]. The MUC1/Y protein, which lacks the tandem repeats, exhibits features common to cytokine receptors and may be involved in signal transduction [5,8,10]. The secreted MUC1/SEC isoform may act as a ligand for the MUC1/Y membrane receptor protein [8]. Recently, it has been shown that the MUC1/SEC protein may also inhibit proliferation of T-cells [11].

It appears that expression of the various MUC1 protein isoforms, especially those involved in signal transduction, must be highly regulated. It was shown that expression of the MUC1 gene is controlled at the transcriptional level [2,12–14]. Nonetheless, knowledge about regulation of its expression is limited. Although several transcription *cis*-elements have been identified in the MUC1 promoter [12–14], these data are not sufficient to build a detailed map of the *cis*-elements present in the promoter. On the other hand, such a map would allow one to evaluate the potentials of the MUC1 gene to be transcribed in different types of cells. Still undetermined is whether transcription of the MUC1 gene is associated with synthesis of a single primary transcript or whether several primary transcripts are generated at various transcriptional start sites. It is known that the 0.8 kb *Sac*I/*Xmn*I fragment, located at the 3′ end of the MUC1 promoter, is a minimal fragment that can direct transcription of the test genes at maximal levels [14]. However, it has not been studied whether this region is also competent to direct expression of the MUC1 protein.

We addressed these questions in the present study. We performed a computer analysis of the MUC1 promoter DNA sequence for potential transcription *cis*-elements and constructed the map of transcriptional *cis*-elements found in the *Sac*I/*Xmn*I region of the MUC1 promoter. We analyzed functional activities of various MUC1 promoter fragments and showed that the *Sac*I/*Xmn*I fragment controls the expression of at least tandem repeat containing isoforms of the MUC1 protein. We also showed that both the 3′ and 5′ end regions of the *Sac*I/*Xmn*I fragment possess independent promoter activities in transfection assays. Finally, we examined the transcription start site(s) (TSS) of the MUC1 gene and found that in mammary carcinoma T47D cells, transcription of this gene is associated with multiple TSSs and that additional TSSs may be activated by treatment of the cells with TGFβ1.

2. Materials and methods

2.1. Cell cultures

Human breast adenocarcinoma T47D cells, human breast epithelial HBL-100 cells and canine fetal thymus Cf2Th cells were propagated as described [15–17].

2.2. Plasmids

The MUC1 promoter containing vector was constructed by replacing of the HMG promoter in the pHMG/Cl642/MUC1 expression vector [18], which includes a 4.3 kb MUC1 cDNA sequence, by the *Sac*I/*Xmn*I fragment of the 5′ end sequence of the MUC1 gene (Figs. 2 and 3A). A promoter-less pCl642/MUC1 plasmid was constructed by excision of the HMG promoter from the pHMG/C1642/MUC1 expression vector.

The CAT test gene containing constructs were prepared as follows. The linker oligonucleotide including *Hin*dIII, *Bst*XI, *Xma*I, *Sac*I, *Xmn*I, *Bsu*36I and *Xba*I restriction sites was inserted into the cloning region of the promoter-less pCAT/enh plasmid (Promega), enabling incorporation of the *Sac*I/*Xmn*I MUC1 promoter fragment. Digestion of the resulting plasmid, pPr/CAT/enh, with *Bst*XI, *Xma*I or *Bsu*36I

*Corresponding author. Fax: (972) (3) 640 2046.
E-mail: josephz@post.tau.ac.il

PII: S0014-5793(99)01452-0

restriction enzymes produced deletion mutants used in CAT transfection assays.

2.3. Transient transfection assay

Cells were transfected utilizing the Ca-phosphate precipitation method or liposomal technique [19]. After 24–48 h, MUC1 protein expression was ascertained by immunohistochemical staining with H23 monoclonal antibody (mAb) [20] that recognizes an epitope within the MUC1 tandem repeat array. CAT activity in lysates of the transfected cells was detected as described [19].

2.4. RNase protection assay (RPA)

RPAs were performed using Ambion HSH kit according to the manufacturer's recommendations [21]. Total RNA was extracted from the cells [22]. The RNA probe for RPA was synthesized by SP6 RNA polymerase from a *Sac*I/*Xmn*I MUC1 fragment cloned into pGEM-7Zf plasmid and was size-purified in a 6% polyacrylamide gel (Fig. 4, IIC).

2.5. Computer analysis

The computer analysis of the potential transcription *cis*-elements in the MUC1 promoter was performed by using the TRANSFAC and TSSG databases on transcription factors and their DNA binding sites [23,24]. The MUC1 promoter sequence was obtained from the GenBank NCBI (accession no. X69118). The degree of sequence homology considered significant and used in this study for most of the *cis*-elements is 0.8–1.0.

3. Results

3.1. Competence of the SacI/XmnI fragment for MUC1 gene transcription

Previous studies showed that the *Sac*I/*Xmn*I region of the MUC1 promoter could drive maximal transcription of test genes [13,14,25]. In order to study the competence of the *Sac*I/*Xmn*I fragment in controlling expression of MUC1 protein isoform(s), we inserted the *Sac*I/*Xmn*I fragment upstream to the human MUC1 coding sequence, cloned in an expression vector, and analyzed the promoter activity by transient transfection of murine DA3 cells. Transfected cells were screened using the mAb H23 specific for tandem repeat containing isoforms of the MUC1 protein.

Fig. 1A shows that murine cells transfected with pPr/CL642/MUC1 plasmid containing the *Sac*I/*Xmn*I sequence were able to express the human MUC1 protein. The same expression vector but devoid of the *Sac*I/*Xmn*I fragment, pPr/CL642/MUC1, was insufficient for MUC1 protein synthesis in transfected cells (negative control, data not shown). For a positive control, pHMG/CL642/MUC1 was used in which the *Sac*I/*Xmn*I fragment was substituted with the housekeeping HMG promoter. The resulting plasmid expressed MUC1 protein at high levels (Fig. 1B).

Thus, the *Sac*I/*Xmn*I fragment of the MUC1 promoter contains a set of *cis*-elements necessary for expression of the MUC1 protein (at least its tandem repeat containing isoforms) in mouse mammary epithelial cells.

3.2. A computer analysis of potential cis-elements in MUC1 promoter

As the *Sac*I/*Xmn*I fragment can drive transcription of the MUC1 tandem repeat containing isoform(s), we wanted to know which transcriptional *cis*-elements may participate in this control. For this purpose, we have performed a computer search for potential transcription *cis*-elements present in the *Sac*I/*Xmn*I fragment.

Transcription results from interplay between transcription

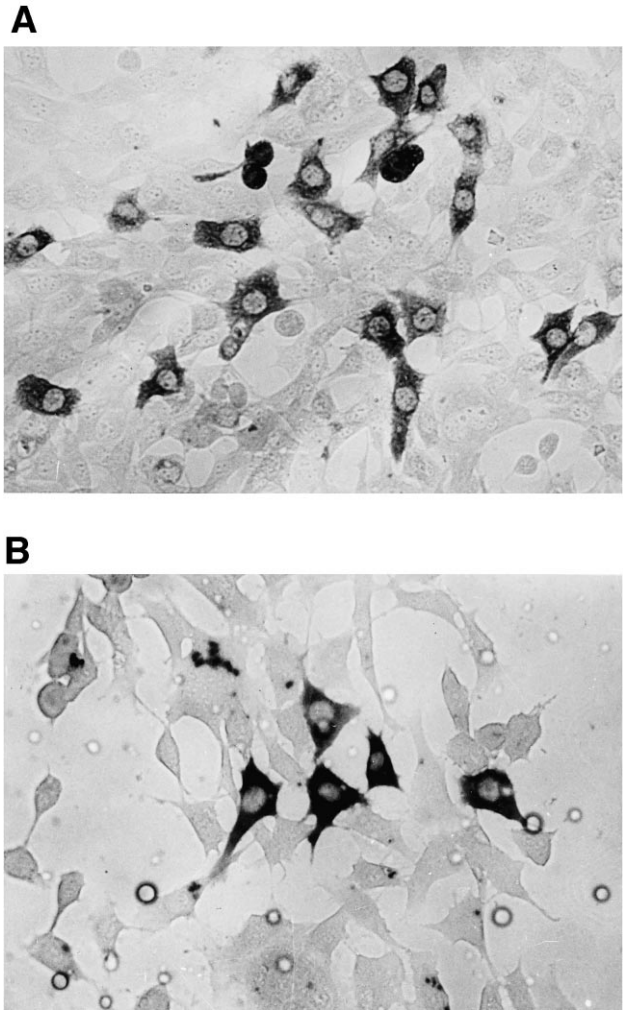


Fig. 1. Human MUC1 protein expression in transiently transfected mouse DA3 cells. A: Immunostaining of the MUC1 protein in cells transfected with plasmid pPr/CL642/MUC1. B: Immunostaining of the MUC1 protein in cells transfected with pHMG/CL642/MUC1 plasmid (positive control). For negative control, pCL642/MUC1 plasmid lacking promoter sequence was used (data not shown). For immunostaining, monoclonal H23 mAb specific for tandem repeats of the MUC1 protein was used.

responsible proteins (transcription factors and RNA polymerase) with transcription specific DNA sequences (*cis*-elements). These two principle components, proteins and DNA sequences, are equal and necessary partners in the transcription process and absence of either one will abolish transcription. A computer analysis of promoter DNA sequences is an appropriate method for identifying transcription *cis*-elements. What transcription factors will interact with the *cis*-elements present in a given promoter depends on the cell type, transcription factor composition, their concentration and affinity for the specific *cis*-elements. The computer analysis cannot answer these questions. However, if the degree of homology between the promoter sequences and the *cis*-element consensus sequences is sufficiently high, it may demonstrate the promoter *cis*-element content, draw the structural complexity of a promoter and reveal the potentials of a given gene to be transcribed in different types of cells. This is the main value of computer *cis*-elements analysis.

Presence of limited numbers of transcription *cis*-elements in the MUC1 promoter has already been reported [13,14,25]. Our search revealed a much more complex picture of the *cis*-element setting. We have discovered 104 different elements in the sense strand and 67 elements in the antisense strand and built a detailed map of the *cis*-elements present in the MUC1 promoter (Fig. 2). Our analysis showed that among these elements, there are (1) ubiquitous transcription *cis*-elements such as TATA-box, CCAAT-, E- and GC-boxes and elements that bind AP1, AP2, AP3, AP4, CTF/NF1, ER, PR, Sp1, STAT1, STAT3, STAT5 and YY1 ubiquitous transcription factors, (2) the elements that regulate transcription in mammary epithelial cells (MAF, MGF, MP4, RME, PMR, SpA and WAP), (3) the elements that are specific for transcription in hematopoetic cells (BKLF/TEFII, GATA1, v-Myb, c-Myb and MZF1), (4) the elements that drive transcription in immunospecific cells (AML-1, Gfi-1, Ikaros, IL-6 RE, LyF1, NF-GMCSF, NF-mu-E1, NF-Y, Pu-box, SRY, TCF-1, TdT Inr, XBP1 and W-element), (5) the elements that are specific for transcription in hepatocytes (ARP-1, HNF-5, LF-A1 and H-APF1), (6) the elements that control transcription in muscle cells (Myo D, Nkx-2.5 and SEF1) and (7) the elements specific for viral promoters (JCV, LBP-1, LVC, PEA1 and PEA3, PV-E2, T-Ag SV40, retroviral TATA-box, TEF-1, TEF-2 and TFII-ML-Inr2) (Fig. 2) [26–30].

As one can see in Fig. 2, the MUC1 promoter is characterized by a high density of transcriptional *cis*-elements distributed along the promoter DNA. Extensive overlapping sequences specific for different *cis*-elements create clusters of *cis*-elements. Some clusters are composed primarily of *cis*-elements specific for a definite transcription factor (for example, the cluster located at −60/−45 is specific for Sp1 factor). Other clusters contain overlapping sequences specific for several functionally unrelated factors (for example, the cluster located at −516/−474 nucleotides is composed of binding sites specific for Ikaros proteins, STAT, ELK-1, c-REL, MZF-1, IL-6, MGF, Nkx-2.5, NF-kB, c-Myb and RME). One may speculate that specific combination of different *cis*-elements in a cluster and the position of a cluster along the promoter DNA sequences may determine the mode of transcription of a particular gene in a definite type of cells. In different cells, different clusters of the same promoter may be in contact with or be in close proximity to the initiation transcription complex, thus specifying the pattern of transcription.

Among different *cis*-element combinations found in the MUC1 promoter, the combinations of the overlapping *cis*-elements which may interact with the transcriptional activators or repressors are of particular interest. For example, the combinations of YY1 repressor binding sites with *cis*-elements specific for initiator TFII-ML-Inr-2 (−669/−651) or for various *trans*-activators such as MGF (−424/−412), Sp1 (−424/−396), MAF (−424/−392), HNF-5 (−314/−302), PEA1 (−159/−140), PV-E2 (−159/−137) and c-Myb (−159/−131) are abundant in the MUC1 promoter. Another example is the NF-kB activator specific sequence (−373/−364) which overlaps the sequence specific for δEF1 repressor (−366/−355). At last, at the very 3′ end of the MUC1 promoter, there is a binding site for the δEF1 repressor (−4/+7) which overlaps one of the main basal TSSs (Figs. 2 and 4). This combination of the overlapping sequences creates a possibility for inhibition of the MUC1 gene transcription in cells where δEF1 repressor successfully competes for the indicated site with the transcription initiation factors.

In conclusion, the computer analysis of the potential *cis*-elements revealed a high level of structural complexity of the MUC1 promoter that might be reflected in functional complexity as well. To test this possibility, we performed a functional analysis of different parts of the MUC1 promoter.

3.3. Functional analysis of the MUC1 promoter: transcription activities associated with the 5′ and 3′ end regions of the SacI/XmnI fragment

For a more detailed analysis of the *Sac*I/*Xmn*I promoter region, the CAT transfection assays were performed. The *Sac*I/*Xmn*I fragment or its deleted forms were inserted into pCAT/enh vector and the resulting constructs were analyzed by CAT assays in three different cell lines: human breast carcinoma T47D cells, human mammary epithelial HBL-100 cells and canine fetal thymus Cf2Th cells.

The *Bst*XI/*Xmn*I (−593/+31) fragment obtained by deletion of the *Sac*I/*Bst*XI sequence from the full-length *Sac*I/*Xmn*I fragment maintained 74.3 and 77.4% of the maximal activity of the *Sac*I/*Xmn*I fragment in epithelial HBL-100 and T47D cells, respectively (Fig. 3B, line 2). Cf2Th thymus cells, transfected with the *Bst*XI/*Xmn*I fragment, supported only 29.2% of the full-length fragment CAT activity. We assume that some elements deleted from the *Bst*XI/*Xmn*I fragment are more crucial for expression in thymus cells than in mammary epithelial cells. To test this assumption, a synthetic 25 base oligonucleotide corresponding to a part of the cluster XIII located at the −618/−593 position (Fig. 2) was added to the 5′ end of the *Bst*XI/*Xmn*I fragment and the new construct was analyzed by CAT assay. CAT activity was increased 2.2-fold relative to the *Bst*XI/*Xmn*I fragment in thymus Cf2TH cells, whereas in epithelial cells, only a small increase in CAT activity was detected (Fig. 3B, line 3). These results showed that the added 25 bp sequence indeed contains *cis*-elements that are more essential for transcription from the MUC1 promoter in thymus cells than in epithelial cells. The computer analysis indicates that among the *cis*-elements present within this 25 bp sequence, there are four elements (α-IRE, Ikaros 1, Ikaros 2 and Pu-box) (Fig. 2) that are specific for gene expression in immune cells [23,24,29,30].

Further analysis of the MUC1 promoter deletion mutants showed that CAT activity of the construct, containing the *Xma*I/*Xmn*I fragment (−410/+31), decreased to 20% in Cf2Th cells and to 37.8 and 32.8% in T47D and HBL-100 cells, respectively (Fig. 3B, line 4).

Truncated fragments of the MUC1 promoter (*Bst*XI/*Xmn*I and *Xma*I/*Xmn*I) that contained the TATA-box but lacked 5′ upstream sequences exhibited CAT activities at lower levels than the construct containing the full *Sac*I/*Xmn*I fragment. In order to determine whether sequences located 5′ upstream to the TATA-box possess TATA-box independent intrinsic promoter activity or whether they only supply enhancing elements to the TATA-based promoter, we deleted the TATA-box from the *Sac*I/*Xmn*I fragment. The resulting *Sac*I/*Bsu*36I (−724/−298 bp) fragment was tested for promoter activity in a CAT assay (Fig. 2B, line 5). Surprisingly, the *Sac*I/*Bsu*36I 'TATA-less' fragment revealed a very high CAT activity in the Cf2Th system (212.8%) and a quite substantial activity in epithelial cells (80%). Although the S*ac*I/*Bsu*36I fragment does not contain a TATA-box, it includes two other promoter

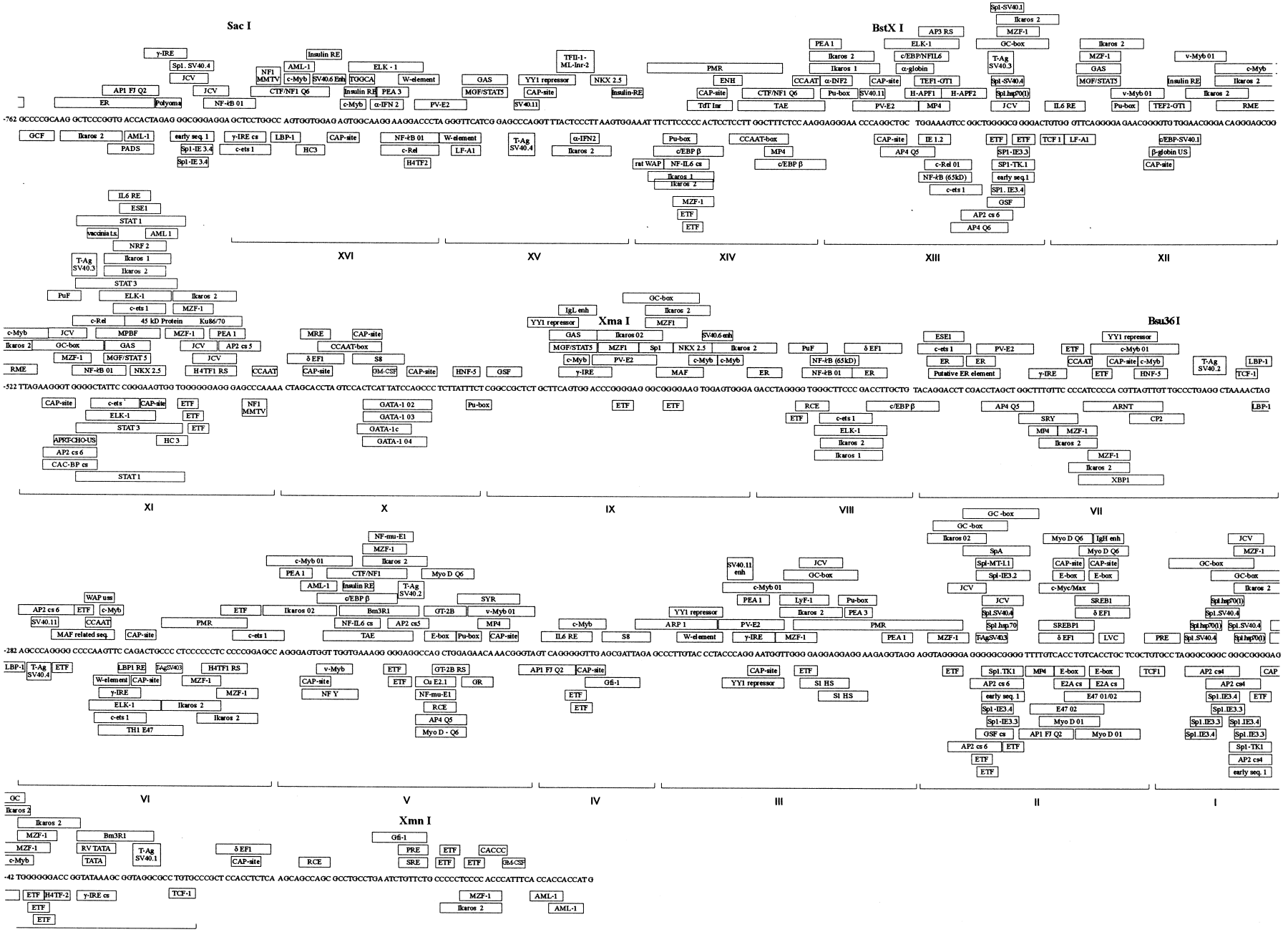

Sac I
BstX I
Xma I
Bsu36 I
Xmn I
XVI
XV
XIV
XIII
XII
XI
X
IX
VIII
VII
VI
V
IV
III
II
I
-762 GCCCCGCAAG GCTCCCGGTG ACCACTAGAG GGCGGGAGGA GCTCCTGGCC AGTGGTGGAG AGTGGCAAGG AAGGACCCTA GGGTTCATCG GAGCCCAGGT TTACTCCCTT AAGTGGAAAT TTCTTCCCCC ACTCCTCCTT GGCTTTCTCC AAGGAGGGAA CCCAGGCTGC TGGAAAGTCC GGCTGGGGCG GGGACTGTGG GTTCAGGGGA GAACGGGGTG TGGAACGGGA CAGGGAGCGG
-522 TTAGAAGGGT GGGGCTATTC CGGGAAGTGG TGGGGGGAGG GAGCCCAAAA CTAGCACCTA GTCCACTCAT TATCCAGCCC TCTTATTTCT CGGCCGCTCT GCTTCAGTGG ACCCGGGGAG GGCGGGGAAG TGGAGTGGGA GACCTAGGGG TGGGCTTCCC GACCTTGCTG TACAGGACCT CGACCTAGCT GGCTTTGTTC CCCATCCCCA CGTTAGTTGT TGCCCTGAGG CTAAAACTAG
-282 AGCCCACGGG CCCCAAGTTC CAGACTGCCC CTCCCCCCTC CCCCGGAGCC AGGGAGTGGT TGGTGAAAGG GGGAGGCCAG CTGGAGAACA AACGGGTAGT CAGGGGGTTG AGCGATTAGA GCCCTTGTAC CCTACCCAGG AATGGTTGGG GAGGAGGAGG AAGAGGTAGG AGGTAGGGGA GGGGGCGGGG TTTTGTCACC TGTCACCTGC TCGCTGTGCC TAGGGCGGGC GGGCGGGGAG
-42 TGGGGGGACC GGTATAAAGC GGTAGGCGCC TGTGCCCGCT CCACCTCTCA AGCAGCCAGC GCCTGCCTGA ATCTGTTCTG CCCCCTCCCC ACCCATTTCA CCACCACCAT G

Fig. 2. The map of the potential transcription *cis*-elements present in the *Sac*I/*Xmn*I fragment of the MUC1 promoter. The *cis*-element abbreviations are as described [12,13,22–27,33–35]. Roman numerals correspond to the clusters located along the *Sac*I/*Xmn*I fragment in 3′ to 5′ direction. ←

elements, transcription initiators, Inr, at −661/−653 and −634/−627 positions which may initiate formation of RNA polymerase II transcription complexes in TATA-box-less promoters [26].

The main conclusion of the functional analysis performed in this study is that both the 5′ and 3′ regions of the *Sac*I/*Xmn*I fragment possess independent promoter activities in transfection assays. On this basis, we hypothesize that the MUC1 gene exhibits characteristics of a 'dual' promoter.

3.4. TSSs of the MUC1 gene in T47D cells

Detection of promoter activities, associated with different regions of the MUC1 promoter, raised the possibility that transcription of the MUC1 gene may occur from different TSSs. To test this possibility, we analyzed transcription of the MUC1 gene in T47D cells using a RPA.

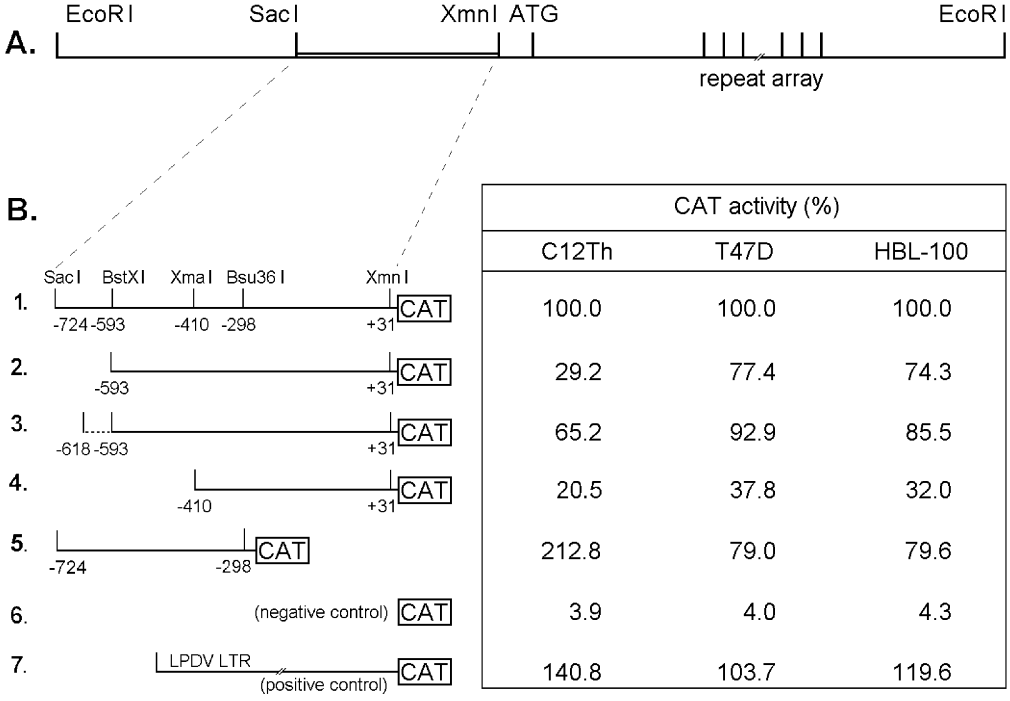
Four TSSs associated with transcription of the MUC1 gene were detected in T47D cells grown under regular growth conditions (Fig. 4, IIA). The major TSS, closest to the ATG start codon, is located 24 nucleotides downstream from the TATA-box (30–35 nucleotides upstream to the *Xmn*I restriction site as indicated in Fig. 4, I) and corresponds to a classical TATA-box-controlled TSS [27,28]. The second TSS virtually overlaps the TATA-box. The third and the fourth TSSs are located 20–25 and 40–45 nucleotides upstream to the TATA-box, respectively, and, therefore, are probably controlled by elements other than TATA-box.

Thus, transcription of the MUC1 gene in T47D cells is associated with multiple transcription start points. It should be pointed out that transcription from the fourth TSS appears to be rather weak and could be observed only after a long exposure (7 days). However, no additional TSSs were detected even after longer (2 weeks) exposures.

3.5. TGFβ1-induced TSSs

As we showed above, in T47D cells cultivated under normal growth conditions, the MUC1 specific TSSs were positioned between +1 and −70 nucleotides at the 3′ end of the promoter (Fig. 4, IIA). No other TSSs located at the 5′ end of the MUC1 promoter were observed in cells cultured under these conditions. On the other hand, the CAT transfection experiments demonstrated promoter activity associated also with the 5′ end −724/−298 fragment. This fragment contains many transcription *cis*-elements (Fig. 2), including two initiators (−661/−653, −634/−627) which might drive independent transcription activity of this fragment. Based on these data, we presumed that additional TSSs might be activated following treatment of the cells with various activators.

The T47D cells possess TGFβ receptors [31] and the *Sac*I/*Xmn*I fragment contains two TGFβ-activated elements at −625/−609 and −224/−208 positions (Fig. 2). Therefore, we chose TGFβ1 as a potential activator of the MUC1 gene in T47D cells. Indeed, TGFβ1 activated two additional TSSs



	CAT activity (%)		
	C12Th	T47D	HBL-100
1.	100.0	100.0	100.0
2.	29.2	77.4	74.3
3.	65.2	92.9	85.5
4.	20.5	37.8	32.0
5.	212.8	79.0	79.6
6.	3.9	4.0	4.3
7.	140.8	103.7	119.6

Fig. 3. CAT assay of the MUC1 promoter activity. A: A scheme of the MUC1 gene. B: CAT activities in T47D, HBL-100 and Cf2Th cells transfected with pPr/CAT/enh plasmid containing the *Sac*1/*Xmn*1 fragment of the MUC1 gene promoter (line 1) or its deletion mutant forms (lines 2–5). For negative control, pCAT/enh plasmid without promoter sequence was used (line 6). For positive control, pCAT/enh plasmid containing LTR of LPD retrovirus was used (line 7).

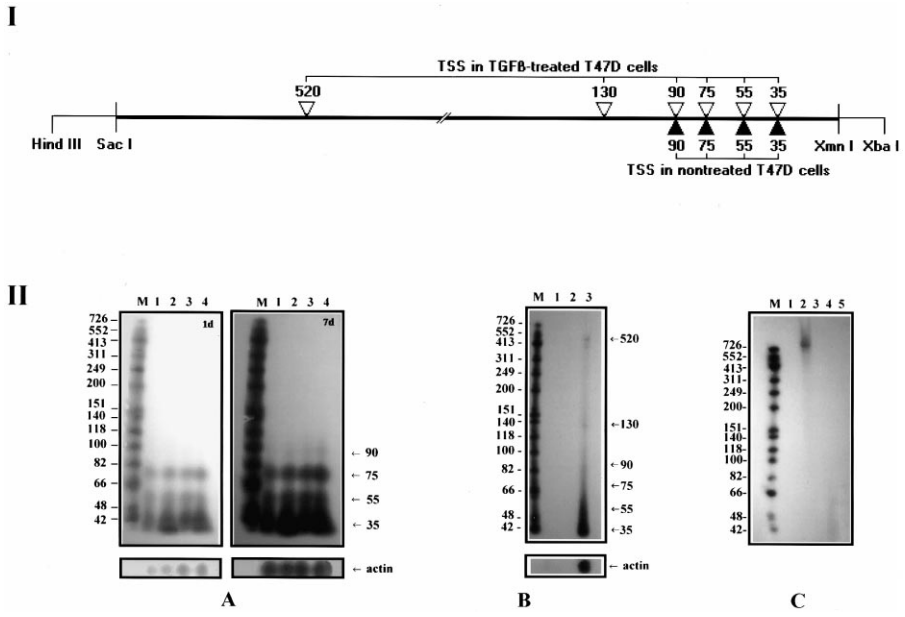


Fig. 4. Detection of the TSSs associated with the MUC1 gene transcription in T47D cells. I: A scheme of the TSS localization in the *Sac*1/*Xmn*1 fragment of the MUC1 gene promoter. The filled triangles indicate the positions of the TSSs associated with the MUC1 gene transcription in non-treated T47D cells, the opened triangles correspond to TSSs in T47D cells treated with TGFβ1. II: Detection of the MUC1 gene specific TSSs by electrophoresis of the RNase-protected fragments in a 6% polyacrylamide gel with 8 M urea. A: RNase-protected fragments of the MUC1 specific transcripts from T47D cells. Lanes 1–4 contain 10, 15, 30 and 50 μg of total RNA, respectively. M: markers. B: RNase-protected fragments of the MUC1 specific transcripts from T47D cells treated with TGFβ1 (5 ng/ml, 48 h). Lanes 1 and 2: empty, lane 3 contains 50 μg RNA. M: markers. C: Detection of the size and specificity of the MUC1 promoter (*Sac*I/*Xmn*I specific) RNA probe. Lanes 1, 3 and 5 are empty. Lane 2: RNA probe hybridized with 50 μg yeast tRNA and non-treated with RNase. Lane 4: RNA probe hybridized with 50 μg yeast tRNA and treated with RNase. M: markers.

(Fig. 4, IIB) located between −500 and −550 nucleotides and at approximately the −130 nucleotide position.

These data confirmed our results obtained in transfection assays that both the 5′ and 3′ regions of the *Sac*I/*Xmn*I fragment of the MUC1 promoter possess promoter activities and may drive transcription of the MUC1 gene from different TSSs.

4. Discussion

Interaction of transcription factors with specific *cis*-elements is a cornerstone of transcriptional regulation. More complex patterns of gene transcription demand increasingly elaborate promoter structure and functions. The MUC1 gene directs synthesis of several protein isoforms with a diverse spectrum of functions [4–9]. Although expression of the MUC1 gene is observed mostly in epithelial cells, its transcription, albeit at low levels, was detected also in other types of cells: normal and neoplastic plasma cells [32,33], follicular dendritic cells [32], Hodgkin and non-Hodgkin lymphomas cells [34,35], myofibroblasts and perineurial cells [32]. Recently, it has been shown that MUC1 gene expression can also be induced in T-cells [11]. The MUC1 gene expression is not only cell and tissue specific, but also developmentally regulated [36]. Alltogether, these data allow one to suppose that the MUC1 promoter should have a high level of complexity. As the MUC1 promoter can drive transcription in the several types of cells mentioned above, a priori, it should contain different *cis*-elements that are specific for transcription in, at least, these cells.

The results obtained by our analysis of the potential *cis*-elements in the MUC1 promoter are in agreement with this assumption. They revealed extensive structural intricacy of the promoter. The MUC1 promoter contains the transcriptional *cis*-elements specific for mammary epithelial cells, hematopoetic cells, B- and T-lymphocytes, hepatocytes and muscle cells. This means that theoretically, the MUC1 gene has the potential to be transcribed in these cells, which correlates well with recent experimental data demonstrating MUC1 gene expression in cells mentioned above [11,32–35]. Thus, the MUC1 promoter seems to be unique among eukaryotic promoters in its assortment of diverse regulatory elements. It includes elements characteristic of differentiation specific, facultative and housekeeping genes, as well as viral genes. Multiple, overlapping *cis*-elements characterized by diverse transcriptional specificity define numerous clusters along the MUC1 promoter DNA. The cluster structure found in the MUC1 promoter suggests 'modules' of regulatory units whose activities may be regulated by spatially and temporally coordinated programs [37].

Structural complexity of the MUC1 promoter correlates with its functional activities revealed by transfection analyses and by RPAs. Our study showed that the 0.8 kb *Sac*I/*Xmn*I fragment, although it comprises only one fourth of the full MUC1 promoter length, contains a set of the *cis*-elements that are competent and sufficient for driving expression of the MUC1 protein, at least its tandem repeat containing isoform(s), in mouse epithelial cells. It is important that this small fragment of the MUC1 promoter controls not only epithelial cell specific transcription but can also regulate gene transcription in lymphoid cells. We showed that the region corresponding to a 25 nucleotide sequence located 5′ upstream to the *Bst*X1 site and composed of T-cell specific *cis*-elements (Pu-box, Ikaros 1, Ikaros 2 and α-IRE) substan-

tially and specifically increased transcription of the CAT test gene in embryonal thymus Cf2Th cells. The presence in the MUC1 promoter of *cis*-elements that may upregulate gene transcription in lymphoid cells correlates well with the observation that the MUC1 gene can be activated in T-cells [11].

Utilizing deleted forms of the *Sac*I/*Xmn*I promoter fragment in CAT transfection assays, we showed that 5′ and 3′ regions of the MUC1 promoter possess independent promoter activities which could be observed when two regions were separated from each other. We hypothesize that TATA-box and GC-boxes located at −30/−25 and −90/−137, respectively, could govern formation of the initial transcription complex (ITC) in the 3′ end-regulated transcription while two initiator elements, TFII-1-ML-Inr2 (−661/−653) and TdT Inr (−634/−627), might control ITC formation in the 5′ end-regulated transcription, as it was observed in some other genes [26].

This hypothesis was confirmed by RPAs which showed that transcription of the MUC1 gene in T47D cells was associated with multiple TSSs located both at the 3′ and 5′ ends of the *Sac*I/*Xmn*I region of the MUC1 promoter. In non-treated cells, the TSSs were generated by the 3′ end of the promoter and were located in the vicinity of the TATA- and GC-boxes. Treatment with TGFβ1 induced additional TSSs located at the 5′ end of the *Sac*I/*Xmn*I fragment.

Previous studies [25] showed that GC-boxes located at −99/−90 bp are involved in MUC1 gene transcription regulation in epithelial cells. We suggest that some of the GC-boxes positioned at the 3′ end of the MUC1 promoter may drive transcription independently of the TATA-box. This may explain our finding of TSSs located upstream to the TATA-box but downstream to the GC-boxes. In this scenario, multiple promoter activities are associated with TATA-box, GC-boxes and Inr elements bestowing MUC1 promoter with characteristics of a mixed polypotent promoter. Further analysis using site specific mutagenesis is currently undertaken to clarify the roles of the discussed elements in transcriptional regulation of the MUC1 gene.

Acknowledgements: We thank Dr M. Hareuveni for providing us with the pHMG/CL642/MUC1 plasmid, Dr I. Tsarfaty for pUN121 plasmid containing the MUC1 gene sequence, Prof. A. Yaniv and Prof. A. Gazit in whose laboratory the CAT assays were performed and Dr M. Gorivodsy for help in computer analysis. This research was supported by the Israel Cancer Association, the Simco Chair for Breast Cancer Research, the Federico Fund for Tel Aviv University and the Barbara Friedman Fund (I.K.) and Susan Komen Breast Cancer Foundation (D.H.W.).

References

[1] Patton, S., Gendler, S. and Spicer, A. (1995) Biochim. Biophys. Acta 1241, 407–424.
[2] Zaretsky, J., Weiss, M., Tsarfaty, I., Hareuveni, M., Wreschner, D.H. and Keydar, I. (1990) FEBS Lett. 265, 46–50.
[3] Abe, M. and Kufe, D. (1990) J. Cell. Physiol. 143, 226–231.
[4] Wreschner, D.H., Hareuveni, M., Tsarfaty, I., Smorodinsky, N., Horev, J., Zaretsky, J., Kotkes, P., Weiss, M., Lathe, R., Dion, A. and Keydar, I. (1990) Eur. J. Biochem. 189, 463–473.
[5] Zrihan-Licht, S., Baruch, A., Elroy-Stein, O., Keydar, I. and Wreschner, D.H. (1994) FEBS Lett. 356, 130–136.
[6] Zhrihan-Licht, S., Vos, H., Baruch, A., Elroy-Stein, O., Sagiv, D., Keydar, I., Hilkens, J. and Wreschner, D.H. (1994) Eur. J. Biochem. 224, 787–795.
[7] Baruch, A., Hartmann, M., Zrihan-Licht, S., Greenstein, S., Burstein, M., Keydar, I., Weiss, M., Smorodinsky, N. and Wreschner, D.H. (1997) Int. J. Cancer 71, 741–749.
[8] Baruch, A., Hartmann, M., Yoeli, M., Adereth, Y., Greenstein, S., Stadler, Y., Skornik, Y., Zaretsky, J., Smorodinsky, N., Keydar, I. and Wreschner, D.H. (1999) Cancer Res. 59, 1552–1561.
[9] Ligtenberg, M., Buijs, F., H.Vos, H. and Hilkens, J. (1992) Cancer Res. 52, 2318–2324.
[10] Pandey, P., Kharbanda, S. and Kufe, D. (1995) Cancer Res. 55, 4000–4003.
[11] Agrawal, B., Krantz, M., Parker, J. and Longenecker, B.M. (1998) Cancer Res. 58, 4079–4081.
[12] Abe, M. and Kufe, D. (1984) Cancer Res. 44, 4574–4577.
[13] Abe, M. and Kufe, D. (1993) Proc. Natl. Acad. Sci. USA 90, 226–231.
[14] Kovarik, A., Peat, N., Wilson, D., Gendler, S. and Taylor-Papadimitriou, J. (1993) J. Biol. Chem. 268, 9917–9926.
[15] Segev, N., Hizi, A., Kerenberg, F. and Keydar, I. (1985) Proc. Natl. Acad. Sci. USA 82, 1531–1535.
[16] Renaud, F., El Yazidi, I., Boilly-Marer, J., Courtois, Y. and Laurent, M. (1996) Biochem. Biophys. Res. Commun. 219, 679–685.
[17] Renshaw, R., Gonda, M.A. and Casey, J.W. (1991) Gene 105, 179–184.
[18] Gautier, C., Mehtali, M. and Lathe, R. (1989) Nucleic Acids Res. 17, 83–89.
[19] Gak, E., Yaniv, A., Sherman, L., Ianconescu, M., Tronick, S. and Gazit, A. (1991) Gene 99, 157–162.
[20] Keydar, I., Chou, C., Hareuveni, M., Tsarfaty, I., Sahar, E., Selzer, G., Chaitchik, S. and Hizi, A. (1989) Proc. Natl. Acad. Sci. USA 86, 1362–1366.
[21] Ambion: Catalog (1996).
[22] Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.
[23] Solovyev, V. and Salamov, A. (1997) in: Proceedings of the International Conference on Intelligent Systems for Molecular Biology (Gaasterland, T., Karp, P., Karplus, K., Ouzounis, C., Sander, C. and Valencia, A., Eds.), pp. 294–302.
[24] Wingender, E., Dietze, P., Karas, H. and Knueppel, R. (1996) Nucleic Acids Res. 24, 238–241.
[25] Kovarik, A., Lu, P., Pear, N., Morris, J. and Taylor-Papadimitriou, J. (1996) J. Biol. Chem. 271, 19140–19147.
[26] Weis, L. and Reinberg, D. (1992) FASEB J. 6, 3300–3309.
[27] Bucher, P. (1990) J. Mol. Biol. 212, 563–578.
[28] Bucher, P. and Trifonov, E. (1986) Nucleic Acids Res. 14, 10009–10026.
[29] Faisst, S. and Meyer, S. (1992) Nucleic Acids Res. 20, 3–26.
[30] Turpaev, K. and Vassetzky, Y. (1990) Genetika 26, 804–816.
[31] Jacowlew, S., Moody, T. and Mariano, J. (1997) Anticancer Res. 17, 1849–1860.
[32] Kamoshida, S. and Tsutsumi, Y. (1998) Pathol. Int. 48, 776–785.
[33] Delsol, G., Al Saati, T., Gatter, K., Gerdes, J., Schwarting, R., Caveriviere, P., Rigal-Huguet, F., Robert, A., Stein, H. and Mason, D. (1988) Am. J. Phatol. 130, 59–70.
[34] Delsol, G., Gatter, K., Stein, H., Erbner, W., Pulford, K., Zinne, K. and Mason, D. (1984) Lancet ii, 1124–1129.
[35] Hall, P., d'Ardenne, A. and Stansfeld, A. (1988) Histopathology 13, 149–160.
[36] Parry, G., Li, J., Stubbs, J., Bissell, M., Schmidhauser, C., Spicer, A. and Gendler, S. (1992) J. Cell. Sci. 101, 191–199.
[37] Yuh, C.-H., Bolouri, H. and Davidson, E. (1998) Science 279, 1896–1902.