ELSEVIER

Journal of Chromatography A, 976 (2002) 95–101

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Purification of proteins specifically binding human endogenous retrovirus K long terminal repeat by affinity elution chromatography

D.O. Trubetskoy*, L.L. Zavalova, S.B. Akopov, L.G. Nikolaev

Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Miklukho-Maklaya 16/10, 117997 Moscow, Russia

Abstract

A novel affinity elution procedure for purification of DNA-binding proteins was developed and employed to purify to near homogeneity the proteins recognizing a 21 base pair sequence within the long terminal repeat of human endogenous retroviruses K. The approach involves loading the initial protein mixture on a heparin–agarose column and elution of protein(s) of interest with a solution of double-stranded oligonucleotide containing binding sites of the protein(s). The affinity elution has several advantages over conventional DNA-affinity chromatography: (i) it is easier and faster, permitting to isolate proteins in a 1 day–one stage procedure; (ii) yield of a target protein is severalfold higher than that in DNA-affinity chromatography; (iii) it is not necessary to prepare a special affinity support for each factor to be isolated. The affinity elution could be a useful alternative to conventional DNA-affinity chromatography.

Keywords: Human endogenous retrovirus; Proteins; DNA

1. Introduction

Many important cellular processes, such as transcription, replication, and recombination, are regulated through specific binding of certain proteins to corresponding *cis*-acting DNA sequences. In particular, sequence-specific DNA-binding proteins are directly involved in the regulation of transcription initiation both in prokaryotes [1] and eukaryotes [2–5]. To study the biochemical properties of these transcription factors, it is necessary to purify the proteins to homogeneity. This enables to characterize the factors and strongly facilitates the raising of antibodies and isolation of the genes encoding these regulatory proteins. However, being regulatory factors, these proteins constitute a very small fraction of cellular proteins, sometimes as small as several molecules per cell. Therefore the purification of these proteins is usually rather difficult and can be achieved only by affinity methods. A great majority of DNA-binding proteins characterized to date was isolated by DNA-affinity chromatography [6]. This method involves coupling synthetic double-stranded oligonucleotide containing a binding site of the protein(s) of interest to agarose support, loading this affinity resin with a partially purified mixture of nuclear proteins and subsequent elution of protein(s) of interest with high ionic strength buffer. Despite successful application of this method to tens of different regulatory proteins, it has several disadvantages. Firstly, the method is quite laborious and time-consuming. It is necessary to prepare affinity supports specific for each protein factor to be isolated. The specificity of the columns is often not

*Corresponding author. Fax: +7-95-330-6538.
E-mail address: dmitrii@humgen.siobc.ras.ru (D.O. Trubetskoy).

0021-9673/02/$ – see front matter
PII: S0021-9673(02)01236-0

sufficient to purify the protein in one cycle, that necessitates rechromatography.

We propose an alternative approach that allows one to avoid some of the above mentioned disadvantages. The approach was applied to purification of the factors capable of binding DNA sequences within solitary long terminal repeats (LTRs) of human endogenous retroviruses. These elements are scattered in several thousand copies throughout the human genome, and they are potentially capable of affecting many cellular processes (for review see Refs. [7–10]). The LTR sequences of one of the most abundant HERV families (HERV-K) harbor a specific binding site for a complex consisting of at least three proteins, the site being located within the 5′ region of the LTR U3 element [11]. Here we report the purification of the proteins which bind this element.

2. Experimental

2.1. Chemicals and equipment used

Phosphate buffered saline, cell culture media and fetal calf serum were obtained from GibcoBRL Life Technologies (Invitrogen, Carlsbad, CA, USA). Heparin–agarose (type II-S), 4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid (HEPES), spermine, spermidine, dithiothreitol, glycerol, phenylmethylsulfonyl fluoride, Brilliant blue R (all reagent grade) were purchased from Sigma (St. Louis, MO, USA). Gel electrophoresis reagents (Tris, acrylamide, *N*,*N*′-methylene bisacrylamide, glycine, sodium dodecylsulfate (SDS)) and reagents for silver staining of the gels were from Bio-Rad (Hercules, CA, USA). Other chemicals (reagent grade or higher) were from Chimmed (Moscow, Russia).

Oligonucleotide primers were synthesized using an ASM-102U DNA synthesizer (Biosset, Novosibirsk, Russia), gels were quantitated with the help of GDS7600 gel analysis system (Ultra-Violet Products, Cambridge, UK) and the Gel-Pro analysis software (Media Cybernetics, Silver Spring, MD, USA).

2.2. Nuclear extract preparation

Nuclear extract (mixture of nuclear proteins) was prepared from mouse Krebs-II ascite carcinoma cells. Each BALB/C male mouse was inoculated with 200 μl of ascite fluid and ascite tumor cells grew in mice for 1 week. The cells were collected by sedimentation in a clinical centrifuge for 5 min at 800 *g*, washed with cold phosphate-buffered saline (PBS), and suspended in 10 m*M* HEPES–KOH, pH 7.9, 25 m*M* KCl, 0.15 m*M* spermine, 0.5 spermidine, 1 m*M* EDTA, 0.5 m*M* dithiothreitol (DTT), 10% glycerol buffer (10 ml per 1 g of the pelleted cells). Nuclear extract was then prepared as described [12] and stored at −70 °C. Human lymphoblast (Jurkat) and HeLa cells were grown in RPMI-1640 and DMEM media, supplied with 10% of fetal calf serum. Nuclear extracts from cultured cells were prepared in the same way as for Krebs-II cells.

2.3. Electrophoretic mobility shift assay and UV-crosslinking

Two complementary oligonucleotides designed to produce two-nucleotide sticky ends (AGGATCGGGGCGGGGCGATC and CTGATCGCCCCGCCCCGATC) were annealed, labeled by filling-in the sticky ends with Klenow enzyme and purified by polyacrylamide gel-electrophoresis as described previously [13]. Electrophoretic mobility shift assay (EMSA), UV-crosslinking and two-dimensional electrophoresis were done using this oligonucleotide as described earlier [11,13]. The relative amount of DNA-binding proteins was estimated using gel images obtained by a charge-coupled device camera system using the Gel-Pro analysis software (Media Cybernetics).

2.4. Fractionation of nuclear extract on heparin–agarose

All operations, unless otherwise indicated, were performed on ice or at 4 °C; 0.4 ml of water was added to 0.8 ml of the nuclear extract with the protein concentration adjusted to about 2.5 mg/ml, and the diluted extract was loaded onto a 0.5 ml heparin–agarose (Sigma) column equilibrated with buffer B (12 m*M* HEPES–KOH pH 7.9, 60 m*M* KCl, 0.12 m*M* EDTA, 0.3 m*M* DTT, 0.3 m*M* PMSF, 12% glycerol). The loading was done at ~0.1 ml/min, the flow-through fraction was collected and

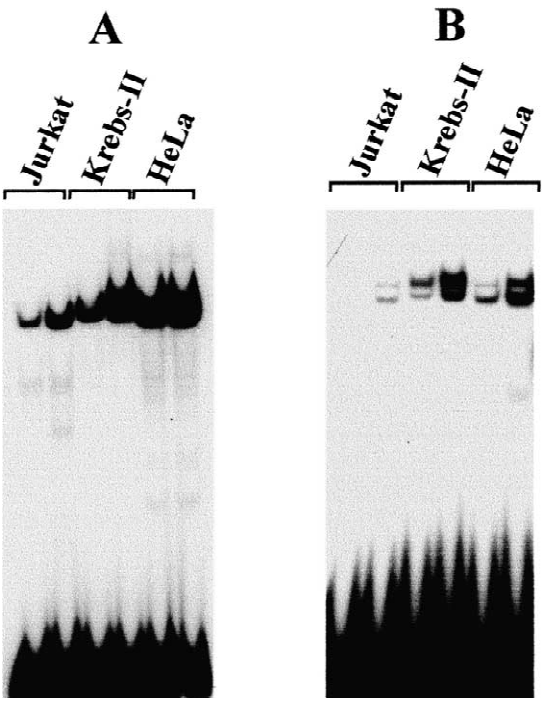

Fig. 1. EMSA using nuclear extracts from Jurkat, Krebs-II, and HeLa cells. (A) 28 bp [11], and (B) 21 bp double stranded oligonucleotide containing the ERFs binding site; 5 and 10 μg of the nuclear extract protein were used for complex formation in the adjacent lanes.

reloaded on the column. The process was repeated five times. The column was then washed with five volumes of buffer B and additionally with five volumes of buffer B containing 200 m*M* KCl. The heparin-bound proteins were eluted with buffer B containing 0.5 *M* KCl. Protein was collected in a 3–4-ml fraction and stored at −70 °C.

2.5. DNA-affinity chromatography

Sepharose with immobilized double-stranded oligonucleotide was prepared as in Kadonaga et al. [6] using the oligonucleotides described in Section 2.3. For fractionation, 3–4 ml of the heparin–agarose purified nuclear extract was diluted with buffer B (without KCl) to a KCl concentration of 60 m*M* and loaded at 4 °C on a 100 μl oligonucleotide–Sepharose column at ~0.1 ml/min, and the flow-through fraction was five times reloaded. The column was further washed with five column volumes of buffer B supplied with 0.1 *M* KCl. Proteins were eluted in 500 μl-fractions with buffer B containing 0.4 *M* KCl. Trichloroacetic acid (TCA) was added to each fraction up to a concentration of 12%, the protein was precipitated for 30 min at 4 °C, pelleted in a microcentrifuge (10,000 *g*, 4 °C, 15 min), washed with 12% TCA and then with acetone, dried on air and dissolved in 20 μl of 20 m*M* HEPES–KOH, pH 7.9. Prior to SDS-gel electrophoresis each 20 μl portion was supplied with 5 μl of SDS loading buffer (0.25 *M* Tris–HCl, 5% SDS, 50% glycerol, 5% 2-mercaptoethanol, 0.025% bromophenol blue) and incubated for 5 min at 100 °C.

2.6. Affinity elution

For affinity elution, the nuclear extract proteins were loaded on the heparin–agarose column and washed with buffer B containing 0.2 *M* KCl exactly as described in Section 2.4. The LTR-binding proteins were eluted with buffer B additionally supplied with the double stranded oligonucleotide containing binding site of the protein complex (Section 2.3) at a final concentration of 30 n*M*. The fractions were collected, protein precipitated with TCA, washed and dissolved for electrophoresis as described in the previous section.

2.7. Polyacrylamide gel electrophoresis (PAGE)

SDS–PAGE was performed as described earlier [14]. Proteins were stained with Brilliant blue R, or with silver nitrate by the method of Ref. [15]. To estimate the protein composition, the protein fractions obtained after each purification step were resolved by 12% SDS–PAGE, stained and quantitated using the Gel-Pro analysis software (Media Cybernetics) and bovine serum albumin resolved in the same gel as a reference.

3. Results and discussion

3.1. Binding site and species specificity of ERFs

Solitary LTRs of human endogenous retroviruses, scattered throughout the human genome, are potentially capable of affecting the expression of closely

located genes. To assess their regulatory potential, we used mobility shift and UV-crosslinking assays to examine the sequence of an LTR belonging to one of the most abundant HERV families (HERV-K) for the presence of binding sites for the host cell nuclear factors [11]. It was found that the LTR sequence harbors a specific binding site for a complex consisting of at least three proteins, ERF1, ERF2 and ERF3. According to the DNase footprinting data, the sequence protected by the complex from DNase digestion was about 28 base pairs (bp) in length and located within the 5′-region of the LTR U3 element. Interestingly, similar proteins can be found not only in human, but also in rodent cells (Fig. 1). Rodent cells do not contain any HERVs, this finding thus indicates that this protein complex can also bind other regulatory sequences.

To characterize the binding site in more detail, we synthesized a set of overlapping oligonucleotides representing different parts of this sequence and used EMSA to test their ability to form complexes with nuclear extract proteins. The shorter, 21 bp double-stranded oligonucleotide, was shown to retain the ability to bind all three protein factors (Fig. 1B). However, in contrast to the full-size 28 bp binding site, it formed two different complexes, one consisting of all ERFs, and the other of only ERF1 and ERF2 (Fig. 2B). Relative intensities of the corresponding two bands varied depending on cell type (Fig. 1B), amount of nuclear extract protein used in EMSA, and other unknown factors. Nevertheless, the protein composition of the complex remains constant (see below).

3.2. Purification of ERF complex using DNA-affinity chromatography

In order to characterize the ERF proteins and isolate their genes, we tried to purify the factors to near homogeneity in quantities sufficient for raising the antibodies and/or determining their amino acid sequence. Firstly, we attempted to use DNA-affinity chromatography [6]. Several cell lines and tissues were tested as a possible source for isolation of these

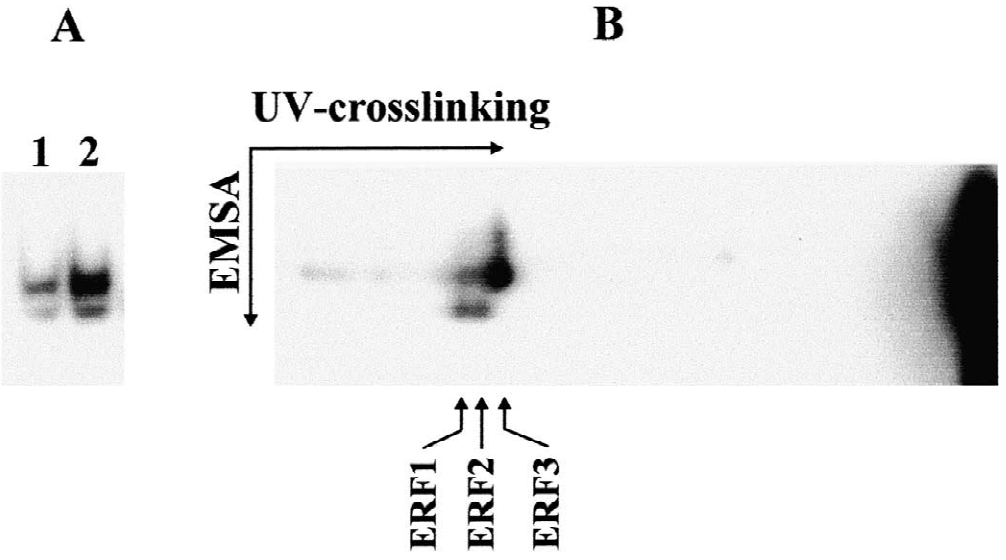

Fig. 2. EMSA using the 21 bp oligonucleotide (A), and protein composition of the EMSA complexes (B) obtained by two-dimensional EMSA–UV-crosslinking–SDS–PAGE analysis. The labeled double-stranded oligonucleotide was incubated with nuclear extract in the presence of excess poly-[dI-dC]/poly-[dI-dC] and resolved by non-denaturing PAGE in the first dimension. The lane containing DNA–protein complexes was cut out and irradiated with UV to form covalently crosslinked DNA–protein complexes. The crosslinked complexes were further separated in the second dimension by denaturing SDS–PAGE [11]. Note that the EMSA complex with lower electrophoretic mobility contains all three ERFs, whereas the complex with higher mobility includes only ERF1 and ERF2.

proteins. The content of ERFs in human cultured cells (Jurkat and HeLa) estimated by EMSA was low, and the amount of cells required for isolation of ~1 μg of each protein was beyond reasonable. On the other hand, preparation of the ERF-containing nuclear extracts from tissues, like human placenta, led to poor yield of nuclei and the nuclear extracts obtained were contaminated with cytoplasmic proteins. The best result was obtained with mouse ascite carcinoma cells which could be easily grown in large quantities and produced high quality nuclear extract. The nuclear extract proteins were obtained in amounts of ~100 mg with relatively high content of the ERF factors as judged by mobility shift assay (not shown).

As a first purification step, we used heparin–agarose binding, widely employed for isolation of DNA-binding factors (e.g. Refs. [16–19]). The negatively charged heparin polymer can be considered as an analog of the DNA phosphate backbone, it is able to bind most of the proteins recognizing DNA sequences. The binding of the nuclear extract to a heparin agarose column led to an about 10-fold enrichment of the eluted protein in the ERF factors (Table 1).

The eluted protein was then loaded onto a DNA-affinity column, eluted with the buffer containing 0.4 *M* KCl, and analyzed by EMSA and SDS–PAGE (Fig. 3). As shown in Fig. 3A, the eluted protein can form a complex with the ERF binding site containing oligonucleotide, characterized by the same mobility as the complex produced with the initial nuclear extract. Fig. 3B demonstrates that the complex contains three proteins with apparent molecular masses of 60,000, 65,000 and 67,000. These values

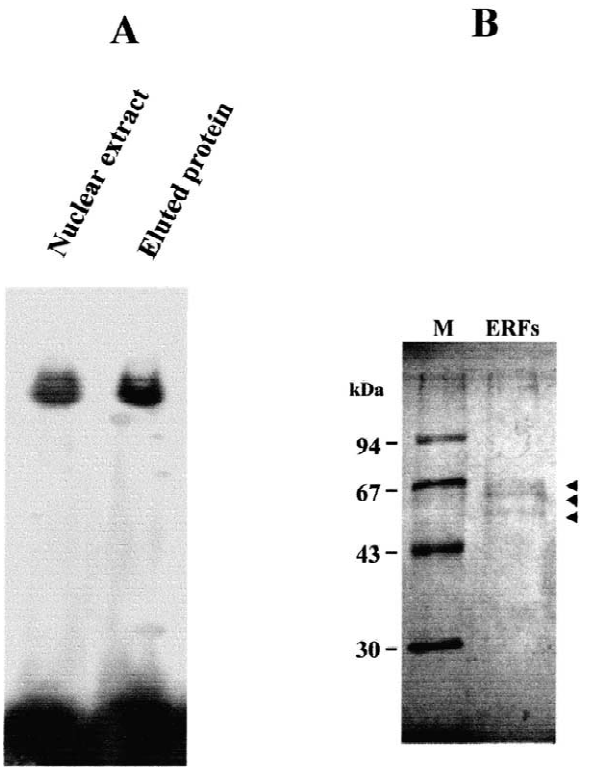

Fig. 3. (A) EMSA using proteins purified by heparin–agarose and DNA-affinity chromatography. (B) Silver-stained SDS–PAGE of the ERF proteins eluted from DNA-affinity column. M, molecular mass marker; kDa, kilodalton.

are somewhat lower than found previously for the ERF factors [11]. However, the earlier data were obtained for proteins crosslinked to oligonucleotide that could lead to some overestimation of the masses.

Although the ERF factors were successfully

Table 1
Purification of DNA-binding proteins using DNA-affinity and affinity elution chromatography

Fraction	Protein (μg)	Protein (%)	Binding activity (%)	Purification (fold)
Initial nuclear extract	3200	100	100	1
Heparin–agarose retained proteins	270	8.5	92	11
Proteins eluted from DNA-affinity column	0.16	0.005	10	2000
Proteins eluted from heparin–agarose by double-stranded oligonucleotide	1.8	0.06	80	1400

purified by DNA-affinity chromatography, their quantities were not sufficient for further analysis. Only 100–200 ng of ERFs could be isolated from 3 mg of the nuclear extract protein, their yield being as low as about 10% (Table 1).

3.3. Affinity elution of ERF complex

The affinity of ERFs to heparin–agarose opened up a possibility of using another approach to purify these proteins. Since the affinity of ERFs for double-stranded oligonucleotide containing the binding site is supposed to be much higher than for heparin, it was reasonable to expect that addition of these oligonucleotides to a low salt elution buffer will lead to elution of ERFs in the form of protein–oligonucleotide complexes. To test this possibility, we absorbed the murine nuclear extract to heparin–agarose similar to the first stage of DNA-affinity chromatography (see above), washed the column with buffer containing 0.2 *M* KCl, eluted the proteins using the same buffer with the ^{32}P-labeled oligonucleotide added, and analyzed the eluted fractions by EMSA (Fig. 4A). It can be seen that the eluted complex has the same electrophoretic mobility as that formed by the proteins of the initial nuclear extract. To assay protein composition of the eluted complex, the fractions eluted as above, but with non-radioactive oligonucleotide, were pooled, concentrated by TCA precipitation, and analyzed by SDS–PAGE (Fig. 4B). The results demonstrated that the protein bands pattern observed after the affinity elution was the same as after the DNA-affinity chromatography and contained three protein bands corresponding to the same apparent molecular masses (cf. Fig. 3B and Fig. 4B). The yield of the ERF proteins in the affinity elution was ~10 times higher than in the DNA-affinity chromatography (Table 1).

4. Conclusions

Using a novel approach to the isolation of DNA-binding proteins, we purified to near homogeneity

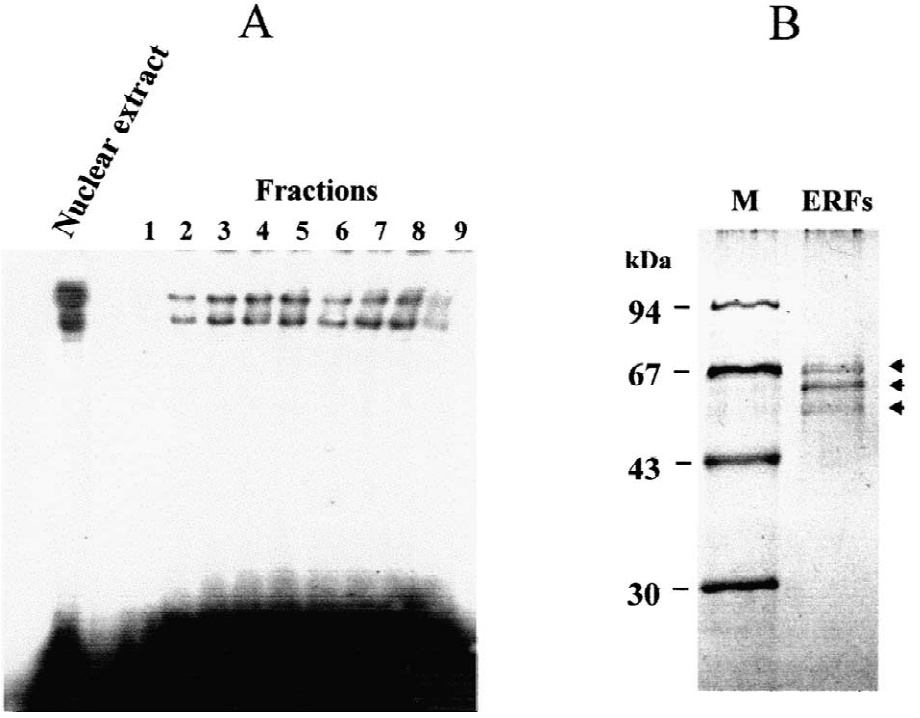

Fig. 4. (A) EMSA of the fractions obtained by affinity elution from the heparin–agarose column with ^{32}P-labeled oligonucleotide containing the ERF binding sites. (B) Brilliant blue R stained SDS–polyacrylamide gel electrophoresis of the pooled fractions 2–8.

(about 1500-fold) the protein factors recognizing a 21 bp sequence within the LTR of human endogenous retroviruses K. The proposed purification method has several advantages over conventional DNA-affinity chromatography.

1. The method is much easier and faster, it is a 1 day–one stage procedure to isolate protein in quantities sufficient for sequencing.
2. The yield of target protein is several times higher than that in DNA-affinity chromatography.
3. There is no need to prepare a special affinity support for each factor to be isolated.

Moreover, it seems possible to fish out proteins with different DNA specificity from one and the same preparation of nuclear extract just by successive elution from the same column with buffers containing different oligonucleotides.

Although the approach described does not permit to isolate factors not retained by heparin–agarose, it can be adapted by using other non-specific supports, e.g. immobilized double stranded polydeoxyinosinic-deoxycytidylic acid (poly-[dI-dC]/poly-[dI-dC]).

Acknowledgements

The authors are grateful to V.K. Potapov and N.V. Skaptsova for oligonucleotide synthesis, to B.O. Glotov for critical reading of the manuscript, and to E.D. Sverdlov for helpful suggestions. This work was supported by grants of the Russian Foundation for Basic Research (01-04-48980 and 02-04-48601).

References

[1] S. Busby, R.H. Ebright, Cell 79 (1994) 743.
[2] T. Boulikas, Crit. Rev. Eukaryotic Gene Expr. 4 (1994) 117.
[3] T.T. Diagana, C.R. Acad. Sci. III 320 (1997) 509.
[4] W.S. Dynan, R. Tjian, Nature 316 (1985) 774.
[5] J.D. Engel, H. Beug, J.H. LaVail, M.W. Zenke, K. Mayo, M.W. Leonard, K.P. Foley, Z. Yang, J.M. Kornhauser, L.J. Ko et al., J. Cell Sci. 16 (Suppl.) (1992) 21.
[6] J.T. Kadonaga, R. Tjian, Proc. Natl. Acad. Sci. USA 83 (1986) 5889.
[7] D.J. Griffiths, Genome Biol. 2 (2001) 1017.
[8] E.D. Sverdlov, FEBS Lett. 428 (1998) 1.
[9] E.D. Sverdlov, Bioessays 22 (2000) 161.
[10] C. Patience, D.A. Wilkinson, R.A. Weiss, Trends Genet. 13 (1997) 116.
[11] S.B. Akopov, L.G. Nikolaev, P.P. Khil, Y.B. Lebedev, E.D. Sverdlov, FEBS Lett. 421 (1998) 229.
[12] J.D. Dignam, R.M. Lebovitz, R.G. Roeder, Nucl. Acids Res. 11 (1983) 1475.
[13] L.G. Nikolaev, Mol. Biol. (Mosk) 30 (1996) 714.
[14] U.K. Laemmli, Nature 227 (1970) 680.
[15] H. Blum, H. Beier, H.J. Gross, Electrophoresis 8 (1987) 93.
[16] E. Genersch, C. Eckerskorn, F. Lottspeich, C. Herzog, K. Kuhn, E. Poschl, EMBO J. 14 (1995) 791.
[17] M.C. Cheng, S.P. Wu, L.F. Chen, S.C. Chen, Planta 203 (1997) 373.
[18] H. Gadgil, H.W. Jarrett, J. Chromatogr. A 848 (1999) 131.
[19] M.C. Estable, M. Hirst, B. Bell, M.V. O'Shaughnessy, I. Sadowski, J. Biomed. Sci. 6 (1999) 320.