Structure of the human transcription factor TFIIF revealed by limited proteolysis with trypsin

Cai Yong[a], Hiromichi Mitsuyasu[b], Zhang Chun[a], Satoru Oshiro[a], Naotaka Hamasaki[b], Shigetaka Kitajima[a,*]

[a] *Department of Biochemical Genetics, Medical Research Institute, Tokyo Medical and Dental University, 1-5-45, Yushima, Bunkyo-ku, Tokyo 113-0034, Japan*

[b] *Department of Clinical Chemistry and Laboratory Medicine, Faculty of Medicine, 3-1-1, Maidashi, Higashi-ku, Fukuoka 812, Japan*

Received 18 June 1998; revised version received 14 August 1998

Abstract In this study, the human general transcription factor IIF (TFIIF), a heteromeric complex of RAP74 and RAP30 subunits, was subjected to limited proteolysis with trypsin. The central region of RAP74 was demonstrated to be highly sensitive to trypsin while both the N- and C-terminal regions contained trypsin-resistant structures. In contrast, RAP30 digestion occurred after proteolysis of RAP74. The digestion pattern of RAP74 recruited into the preinitiation complex showed no marked difference from that of IIF, while RAP30 in the complex was protected from trypsin. These results indicate that RAP74 apparently contains three structural domains, the central one of which is externally surfaced and unstructured, but RAP30 is internally wrapped by RAP74. Furthermore, the accessibility of the central region of RAP74 is unaltered in the minimal preinitiation complex, while RAP30 is involved in promoter recognition through its DNA binding activity.

Key words: Transcription factor IIF; RAP74; RAP30; Limited proteolysis; Structural domain

1. Introduction

The mRNA synthesis of class II genes in vitro is carried out by the transcription machinery, a multiprotein complex composed of general or basal transcription factors including RNA polymerase II. These factors (TFIIB, -D, -E, -F, and -H) that assemble on the minimal promoter sequence of the gene have specific functions and interact with each other to produce a competent initiation complex [1–3]. Among these factors, TFIIF is unique in that it functions in both initiation and elongation [4–7]. In the initiation step, IIF recruits RNA polymerase II into the preformed complex of TBP (TATA-binding protein), IIB, and promoter DNA [8,9]. RAP30 interacts with RAP74 and RNA polymerase II through its N-terminal and central sequences, respectively [10–12]. Its C-terminus exhibits cryptic DNA binding activity [13,14]. RAP74 is also required for reconstitution of both the initiation and elongation activities of IIF [4,5,7,15–17]. Functional analyses of RAP74 have revealed that its N-terminal sequence interacts with RAP30, and is essential and sufficient for supporting the in vitro transcription [11,18,19]. The RAP74 structure deduced from the cDNA sequences of human [15,16], *Xenopus* [20], and *Drosophila* [21,22] contained two highly conserved N- and C-terminal sequences with a long central region rich in acidic and basic charged amino acids. The central region also contains many possible phosphorylation sites [11,23] whose state upregulates transcriptional activity [24]. For *Drosophila*, two more subregions were reported [22]. However, the significance of these regions, especially that of the central charged and C-terminal regions, remains unknown. In this study, TFIIF was expressed as a native hetero-complex of RAP30 and RAP74 in insect cells, and subjected to trypsin digestion in order to determine its native structure in the free form and preinitiation complex. The results showed that the trypsin-sensitive region of RAP74 is clustered in the central domain while both the N- and C-terminal portions contain trypsin-resistant structures. The digestion pattern of RAP74 in the minimal preinitiation complex remained unchanged but RAP30 became more resistant.

2. Materials and methods

2.1. Recombinant human transcription initiation factors

TFIIF was expressed in an insect cell line, SF21 cells, by co-transfection of baculoviruses harboring cDNA encoding each subunit of IIF as described [17]. Approximately 2×10^7 cells in a 175-cm^2 flask were infected with a mixture of 3×10^8 pfu each of BAKPAK6/RAP74 and BAKPAK6/RAP30 baculoviruses [17] in 10 ml of Grace's medium containing 10% fetal bovine serum at room temperature for 1 h. After further incubation at 27°C for 60 h, the cells were harvested and resuspended in 4 ml of ice-cold buffer (20 mM Tris-HCl, pH 7.9, and 5 mM 2-mercaptoethanol) containing 5 μg/ml each of antipain, leupeptin, pepstatin, chymostatin, elastatinal, and phosphoramidon. After the cells had been lysed with a Dounce homogenizer, cell extracts containing expressed IIF were obtained by centrifugation of the homogenates at 35 000 rpm for 60 min. The extracts were then applied to a phosphocellulose (P11) column equilibrated with buffer B (20 mM Tris-HCl, pH 7.9, 0.1 mM EDTA, 10% glycerol, and 5 mM 2-mercaptoethanol) containing 0.3 M NaCl. After washing with the same buffer, IIF was eluted with buffer B/0.6 M NaCl. The fraction containing IIF was diluted to buffer B/0.1 M NaCl by the addition of buffer B, applied to a Q-Sepharose column equilibrated with buffer B/0.1 M NaCl, and then eluted with a linear gradient of NaCl from 0.1 to 0.6 M. IIF was eluted at 0.25 M NaCl and was approximately 95% pure. Human TBP (TATA box-binding protein) and both subunits of IIE that contained a $(histidine)_6$ tag at their N-termini were expressed in *E. coli* strain BL21(DE3) cells, using the pET15b expression vector, and purified by nickel chromatography on a Pro-Bond resin according to the recommended protocol. Recombinant human IIB was obtained as described [25]. Rat liver IIH was kindly provided by Dr. Aso at Tsukuba University, Japan, and human RNA polymerase II was purified from HeLa cells as described [24]. The purities of the factors estimated on SDS-PAGE analysis were approximately 95, 95, 90, 60, and 70% for TBP, IIB, IIE, IIH, and RNA polymerase II, respectively.

2.2. DBPolF or DBPolFEH complex formation

Approximately 10 pmol of each basal transcription factor was used. The amount of TFIIF was substoichiometric as to other basal factors in order to assure complete recruitment of IIF into the complex. Briefly, 0.38 μg TBP and 0.32 μg IIB were first incubated with

*Corresponding author. Fax: (81) (3) 5803-0248.
E-mail: kita.bgen@mri.tmd.ac.jp

PII: S0014-5793(98)01068-0

0.53 μg DNA fragment of the adenovirus major late promoter sequence from −55 to +33 in 60 μl of buffer A (20 mM HEPES-KOH, pH 7.5/0.1 mM EDTA/50 mM KCl/10% glycerol) containing 5 mM $MgCl_2$ and 6 μg poly(dG-dC) at 30°C for 30 min. Then, 1 μg polymerase II (PolII) and 0.34 μg IIF were added in a total volume of 100 μl for the DBPolF complex. A mixture of 0.42 μg IIE and 0.80 μg IIH was further included for DBPolFEH complex formation. The mixture was further incubated for another 30 min for the complex to assemble. Under these conditions, the preinitiation complex formation was detected by gel shift assay as described in [24].

2.3. Trypsin digestion

IIF alone, bound to RNA polymerase II, or recruited into the DBPolF or DBPolFEH complex in buffer A was digested with the indicated amounts of trypsin at 30°C. At each time interval, the reaction was stopped by adding a 15-μl aliquot to SDS-PAGE sample buffer, followed by heating at 95°C. After separation by 12% SDS-PAGE, the RAP74 digestion products were visualized by silver-staining or Western blotting.

2.4. In vitro transcription assay

The in vitro transcription activity was assayed as described [24]. Briefly, the preinitiation complex was assembled by preincubation of pAdML(C2AT) containing the promoter sequence of the adenovirus major late gene and a G-less cassette [24] with TBP, IIB, IIE, IIF, rat IIH, and human RNA polymerase II. Transcription was initiated by adding a mixture of nucleoside triphosphates to a final concentration of 600 μM ATP and UTP, and 25 μM α-^{32}P-CTP (3000 Ci/mmol), followed by further incubation for 60 min. RNA transcripts were analyzed on a 6% polyacrylamide/7 M urea gel, and visualized by autoradiography.

2.5. Other procedures

One mg/ml trypsin (Boehringer Mannheim) was prepared in 1 mM HCl and used after appropriate dilution. The conditions of Western blotting, including electro-transfer to a nitrocellulose membrane, blocking, and detection with chemiluminescent reagent CSPD, were as described [26]. Antibodies against full length RAP74 or RAP30 were prepared as described previously [11], anti-RAP74 (N-16) or anti-RAP74 (C-18) antibodies against the N-terminal 43–58 or C-terminal 498–515 amino acids of RAP74, respectively, were obtained from Santa-Cruze. Anti-TBP and anti-IIB antibodies were also purchased from Santa-Cruze.

3. Results

3.1. Susceptibility of the RAP74 subunit to trypsin digestion

Recombinant IIF was treated with trypsin for various periods and then analyzed by SDS-PAGE. As shown in Fig. 1A, RAP74 was easily digested after short exposure to trypsin and several small digestion fragments were produced (Fig. 1A, lanes 1–7). These fragments were relatively resistant to trypsin since their densities showed no remarkable decrease on further digestion. Among the bands, three major ones of 27, 24, and 22 kDa reacted with both anti-full and anti-N-terminal (43–58) RAP74 antibodies (Fig. 1B, lanes 1–4), demonstrating that these bands represent N-terminal fragments. Two bands of 15 and 13 kDa were detected with both anti-full and anti-C-terminal 498–515 antibodies (Fig. 1B, lanes 1, 2, 5, and 6). When the bands detected on Western blotting were compared with silver-stained proteolytic fragments, it became evident that there were no bands which did not react with anti-N- or anti-C-terminal antibodies. This indicates that the relatively broad region between the N- and C-terminal sequences was digested into very small fragments under these conditions. In fact, no fragments smaller than the N- or C-terminal ones were detected on 17% SDS-PAGE in Tris-Tricine buffer that was employed for smaller peptide analysis (data not shown). In contrast, the density of RAP30 exhibited only a small reduction when non-digested full-sized RAP74 remained. However, its digestion rapidly progressed after completion of RAP74 digestion and produced a 13-kDa fragment. These data indicate that RAP74 contains a trypsin-resistant globular conformation at both its N- and C-termini, but a trypsin-sensitive structure in its central region. RAP30, on the other hand, is protected from trypsin by RAP74 since it was digested after RAP74.

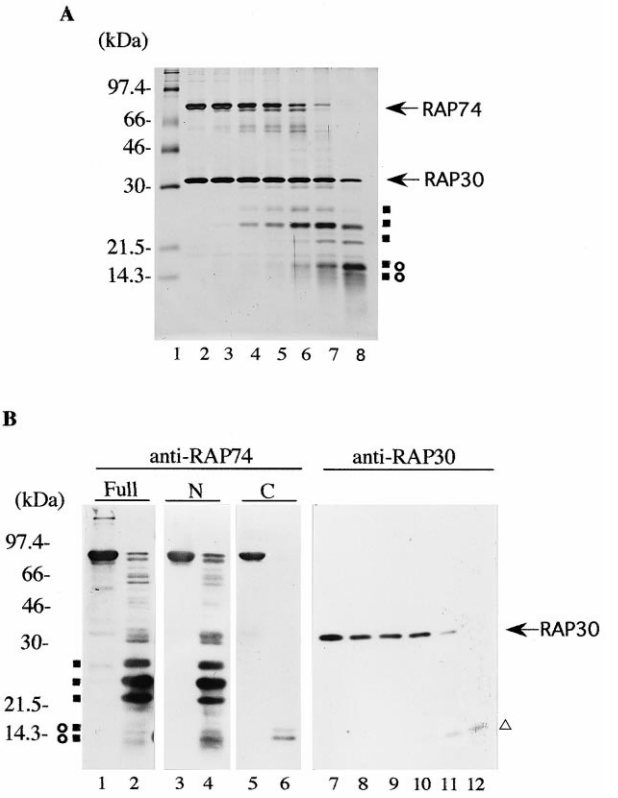

Fig. 1. Eight μg of baculovirus-expressed IIF was subjected to trypsin digestion in a ratio of 1600:1 (IIF to trypsin) at 30°C for the indicated times and the aliquots were analyzed by 11% SDS-PAGE. A: Tryptic fragments were visualized by silver staining. Lane 1: Size marker; lanes 2–8: samples incubated for 0, 1, 3, 5, 10, 20, and 30 min, respectively. Closed squares and open circles indicate major N-terminal and C-terminal fragments, respectively. B: IIF treated with trypsin for 0 (lanes 1, 3, and 5) and 20 min (lanes 2, 4, and 6) were analyzed by Western blotting using anti-full length RAP74 (lanes 1 and 2), anti-N-terminal 43–58 (lanes 3 and 4), or anti-C-terminal 498–515 amino acid antibodies (lanes 5 and 6), respectively. In lanes 7–12, IIF digested for 0, 3, 5, 10, 20, and 30 min was analyzed by Western blotting using anti-RAP30 antibody. The open triangle indicates the 13-kDa major proteolytic fragment of RAP30.

3.2. Limited proteolysis of the IIF-RNA polymerase II complex

TFIIF has been shown to bind to RNA polymerase II (PolII) in solution and stimulate its elongation rate of mRNA synthesis [3–6]. The IIF-PolII complex was formed in the presence or absence of DNA, and treated with trypsin. The digestion profile of RAP74 detected with anti-full RAP74 antibody was essentially the same as that of IIF regardless of the presence or absence of DNA (data not shown).

3.3. Limited proteolysis of IIF in the preinitiation complex

Basal transcription in vitro requires a set of general factors. We assayed the in vitro transcription activity reconstituted

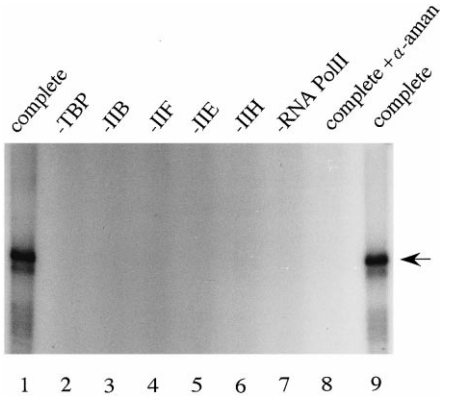

Fig. 2. Reconstitution of the in vitro transcription activity using general transcription factors. Recombinant TBP, IIB, IIF, IIE, and purified IIH were employed for reconstituting the in vitro transcription activity as described in Section 2. Lanes 1 and 9: Complete; lanes 2–7: activity in the absence of TBP, IIB, IIF, IIE, IIH, or RNA polymerase II, respectively; lane 8: complete in the presence of 1 μg/ml α-amanitin. The arrow indicates the band of the correct transcript.

with our preparations of general transcription factors. TBP, IIB, IIF, IIE, IIH, and PolII were all minimal requirements for the activity as shown in Fig. 2. Omission of any factor from the complete reaction mixture totally abolished the transcription activity. This indicated that our factors are capable of supporting the in vitro transcription activity and are not contaminated by each other. Using these factors, the preinitiation complex (PIC) was assembled and then subjected to trypsin digestion. As shown in Fig. 3A, the digestion of RAP74 in the DBPolFEH complex proceeded essentially with the same time course as that for PIC without DNA. Three major N-terminal bands and fragments in the 14-kDa region were protected from digestion, as in the case of IIF, and no specific digestion fragments were produced. Fig. 3B, however, shows that RAP30 digestion was evidently inhibited in PIC with promoter DNA compared with that without DNA. This is likely due to the interaction of RAP30 with promoter DNA. This inhibition of RAP30 digestion was also observed for the DBPolF complex with promoter DNA, but not for that without DNA (data not shown). TBP was also protected from trypsin in PIC with promoter DNA in our assay (data not shown).

4. Discussion

Bacterially expressed RAP30 and RAP74 can associate, but very inefficiently in vitro even after denaturation of the two proteins in urea and subsequent renaturation [4,27]. In contrast, co-expression of RAP30 and RAP74 in insect cells yielded an assembled form of active IIF (Fig. 2 and [17]). In fact, the insect-cell expressed IIF showed the same biochemical properties as native IIF purified from HeLa cells; that is, a stoichiometric heterotetramer of RAP30 and RAP74, the same specific activity of initiation and elongation of transcription, and similar behavior on column chromatography (data not shown). Thus, it is necessary to use well-assembled IIF rather than independently expressed RAP30 or RAP74 for a structural study.

In this study, we found that the central region of RAP74 is highly sensitive to trypsin and externally surfaced, but N- and C-terminal regions contain trypsin-resistant structures. Although the amino acid sequences of the tryptic fragments were not determined, comparison of their sizes on SDS-PAGE with those of RAP74 deletion mutants [11] allowed us to predict that the N- and C-terminal fragments comprise of the N-terminal 180–200 and C-terminal 130–150 amino acids, respectively. Mutagenesis of RAP74 revealed that its N-terminal sequence of 1–172 [18] or 1–205 [11,19,29] interacts with RAP30, and is essential for the ability of in vitro transcription initiation. Our finding that the N-terminal 22–27-kDa region is resistant to trypsin digestion supports the idea that this region has a globular structure, or that it is protected due to the interaction with RAP30. This provides structural evidence for the function of this region of RAP74. RAP74 was also found to interact with IIB [28], PolII and DNA [18] in assays using bacterially-expressed mutants of RAP74 at its C-terminal 358–517, 363–444, and 363–486 sequences, respectively. The present analysis using the in vivo-assembled IIF demonstrated that RAP74 contains another trypsin-resistant

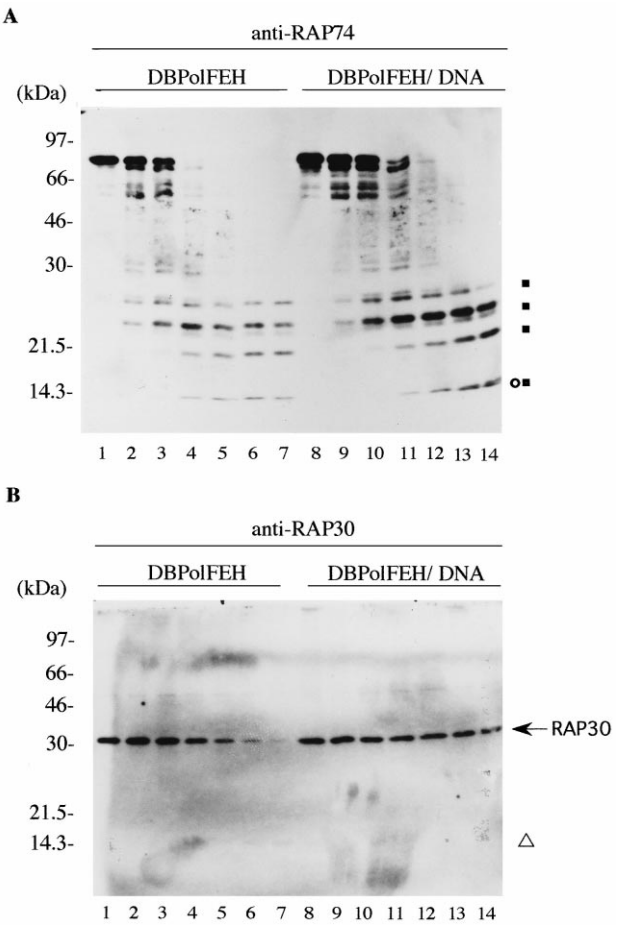

Fig. 3. Trypsin digestion of the in vitro-assembled DBPolFEH complex. The DBPolFEH complex was assembled with (lanes 1–7) or without (lanes 8–14) minimal promoter DNA in vitro as described in Section 2. The complex was then digested with trypsin in a ratio of 200:1 (DBPolFEH/trypsin) for 0 (lanes 1 and 8), 1 (lanes 2 and 9), 3 (lanes 3 and 10), 10 (lanes 4 and 11), 20 (lanes 5 and 12), 30 (lanes 6 and 13), and 50 min (lanes 7 and 14), respectively, separated by 11% SDS-PAGE, and analyzed by Western blotting using anti-RAP74 (A) or anti-RAP30 antibodies (B). The closed squares, open circle in A, and the open triangle in B indicate major N-terminal, C-terminal fragments of RAP74, and the 13-kDa proteolytic fragment of RAP30, respectively.

structure of approximately 130–150 amino acids length at its C-terminus. However, the significance of these interactions was not evaluated in this study since the digestion pattern of RAP74 did not change significantly in the IIF-PolII, DBPolF, or DBPolFEH complex. It is possible that these interactions are neither strong nor stable enough to be assayed by means of trypsin digestion. Alternatively, bacteria-expressed RAP74 might contain a structure different from that of assembled IIF, and exposes cryptic sequences for IIB, PolII, or DNA binding whose function is apparent only in a more complex system than the refined in vitro-assembled assay. In fact, the C-terminus of RAP74 is dispensable for both in vitro initiation and elongation of transcription [18,19,29]. The central region of RAP74, approximately 200–360 amino acids, is considered to be exposed to the outside of the IIF molecule and to have a random structure. This might be due to its notably high content of charged amino acids: several runs or clusters of positive, negative, and mixed amino acids in the 233–252, 302–326, and 180–213 sequences, respectively [23].

RAP30, which is considered to be internally protected by RAP74, became trypsin-resistant in the preinitiation complex containing DNA. It is apparent that this is due to a conformational change of RAP30 evoked by the DNA binding activity at its C-terminus. RAP30 is considered to be involved in the formation of a stable PIC around promoter DNA, in concert with TBP whose digestion by trypsin was also prevented in PIC (data not shown). RAP30 strongly interacted with the coding strand just downstream of TBP at −19 of adenovirus major late promoter DNA [30].

Finally, this study provided some information as to the native structure of TFIIF and might shed light on the structural basis of unique function of IIF in both the initiation and elongation of transcription by RNA polymerase II.

Acknowledgements: We wish to thank Drs. T. Tamura, D. Reinberg, R. Tjian, and J.W. Conaway for the generous gifts of plasmids encoding TBP, IIB, IIE, and IIF, respectively. We also thank M. Tanaka for technical assistance and N.J. Halewood for editing the manuscript. This work was supported in part by Grants-in-Aid from the Ministry of Education, Science, Sports and Culture of Japan, and from the Japan Space Forum to S.K.

References

[1] Conaway, R.C. and Conaway, J.W. (1993) Annu. Rev. Biochem. 62, 161–190.
[2] Zawel, L. and Reinberg, D. (1993) Prog. Nucleic Acids Res. Mol. Biol. 44, 67–108.
[3] Zawel, L. and Reinberg, D. (1995) Annu. Rev. Biochem. 64, 533–561.
[4] Burton, Z.F., Killeen, M., Sopta, M., Ortolan, L.G. and Greenblatt, J. (1988) Mol. Cell. Biol. 8, 1602–1613.
[5] Price, D.H., Sluder, A.E. and Greenleaf, A.L. (1989) Mol. Cell. Biol. 9, 1465–1475.
[6] Flores, O., Ha, I. and Reinberg, D. (1990) J. Biol. Chem. 265, 5629–5634.
[7] Kitajima, S., Tanaka, Y., Kawaguchi, T., Nagaoka, T., Weissman, S.M. and Yasukochi, Y. (1990) Nucleic Acids Res. 18, 4843–4849.
[8] Buratowski, S., Sopta, M., Greenblatt, J. and Sharp, P.A. (1991) Proc. Natl. Acad. Sci. USA 88, 7509–7513.
[9] Flores, O., Lu, H., Killeen, M., Greenblatt, J., Burton, Z.F. and Reinberg, D. (1991) Proc. Natl. Acad. Sci. USA 88, 9999–10003.
[10] McCracken, S. and Greenblatt, J. (1991) Science 253, 900–902.
[11] Yonaha, M., Aso, T., Kobayashi, Y., Vasavada, H., Yasukochi, Y., Weissman, S.M. and Kitajima, S. (1993) Nucleic Acids Res. 21, 273–279.
[12] Tan, S., Conaway, R.C. and Conaway, J.W. (1995) Proc. Natl. Acad. Sci. USA 92, 6042–6046.
[13] Garrett, K.P., Serizawa, H., Hanley, J.P., Bradsher, J.N., Tsuboi, A., Arai, N., Yokota, T., Arai, K., Conaway, R.C. and Conaway, J.W. (1992) J. Biol. Chem. 267, 23942–23949.
[14] Tan, S., Garrett, K.P., Conaway, R.C. and Conaway, J.W. (1994) Proc. Natl. Acad. Sci. USA 91, 9808–9812.
[15] Aso, T., Vasavada, H.A., Kawaguchi, T., Germino, F.J., Ganguly, S., Kitajima, S., Weissman, S.M. and Yasukochi, Y. (1992) Nature 355, 464–466.
[16] Finkelstein, A., Kostrub, C.F., Li, J., Chavez, D.P., Wang, B.Q., Fang, S.M., Greenblatt, J. and Burton, Z.F. (1992) Nature 355, 464–467.
[17] Tan, S., Aso, T., Conaway, R.C. and Conaway, J.W. (1994) J. Biol. Chem. 269, 25684–25691.
[18] Wang, B.Q. and Burton, Z.F. (1995) J. Biol. Chem. 270, 27035–27044.
[19] Kephart, D.D., Wang, B.Q., Burton, Z.F. and Price, D.H. (1994) J. Biol. Chem. 269, 13536–13543.
[20] Gong, D.-W., Hasegawa, S., Wada, K., Roeder, R.G., Nakatani, Y. and Horikoshi, M. (1992) Nucleic Acids Res. 20, 6736.
[21] Kephart, D.D., Price, M.P., Burton, Z.F., Finkelstein, A., Greenblatt, J. and Price, D.H. (1993) Nucleic Acids Res. 21, 1319.
[22] Gong, D.-W., Horikoshi, M. and Nakatani, Y. (1993) Nucleic Acids Res. 21, 1492.
[23] Karlin, S. (1993) Proc. Natl. Acad. Sci. USA 90, 5593–5597.
[24] Kitajima, S., Chibazakura, T., Yonaha, M. and Yasukochi, Y. (1994) J. Biol. Chem. 269, 29970–29977.
[25] Ha, I., Lane, W.S. and Reinberg, D. (1991) Nature 352, 689–695.
[26] Cai, Y., Kitajima, S., Etoh, F., Kinoshita, S., Okubo, K. and Hamasaki, N. (1997) Clin. Exp. Immunol. 109, 488–494.
[27] Wang, B.Q., Kostrub, C.F., Finkelstein, A. and Burton, Z.F. (1993) Protein Exp. Purif. 4, 207–214.
[28] Fang, S.M. and Burton, Z.F. (1996) J. Biol. Chem. 271, 11703–11709.
[29] Joliot, V., Demma, M. and Prywes, R. (1995) Nature 373, 632–635.
[30] Coulombe, B., Joyce, L. and Greenblatt, J. (1994) J. Biol. Chem. 269, 19962–19967.