FEBS Letters 348 (1994) 46–50

FEBS 14199

Transcription factor nuclear factor I proteins form stable homo- and heterodimers

Ulrich Kruse*, Albrecht E. Sippel

Institut für Biologie III/Genetik, Albert-Ludwigs-Universität, Schänzlestr. 1, D-79104 Freiburg, Germany

Received 28 April 1994; revised version received 31 May 1994

Abstract

Nuclear factor I (NFI) proteins constitute a large family of eukaryotic DNA binding proteins. They are involved in viral and cellular aspects of transcriptional regulation and they are capable of stimulating adenovirus initiation of replication. Using in vitro translated NFI proteins encoded by four different chicken NFI genes, we have detected homodimers as well as heterodimers for all combinations tested. The formation of heterodimers was critically dependent on cotranslation, indicating stable dimer formation in the absence of DNA. The unrestricted heterodimerization of NFI proteins adds, beside gene diversity and alternative splicing, another level of diversity to this protein family.

Key words: Transcription factor; DNA binding; Dimerization; Nuclear factor I; TGGCA protein; CCAAT binding transcription factor

1. Introduction

Control of transcription involves the interaction of protein factors with specific DNA sequence elements, that are usually tightly clustered in promoter or enhancer elements [1]. Many of the characterized promoter elements interact with not just one specific binding protein but a family of structurally related proteins. The size of the family can be surprisingly large due to multiple genes and alternative splicing, as is the case for the Fos/Jun family [2] and the nuclear factor I protein family. The diversity of the NFI protein family is generated by four distinct but closely related genes NFI-A, NFI-B, NFI-C/CTF, and NFI-X [3–8], by differential splicing and posttranslational modification by glycosylation [9]. Heterodimer formation between individual members of a protein family can provide an additional mechanism of diversity for increasing the number of DNA-binding proteins [10].

NFI proteins, also sometimes referred to as CCAAT binding transcription factors or CTF [3], bind as dimers to the palindromic consensus sequence 5′-YTGGCA(N)$_3$TGCCAR-3′ [11] functioning as transcription factors [3,12–14] and replication factors [3,15,16]. All NFI proteins have a highly conserved N-terminal domain in common. The N-terminal 220 amino acids of human NFI-C/CTF and the N-terminal 220 amino acids of the rat NF1-L protein have been shown to be sufficient for DNA-binding, dimerization, and stimulation of adenovirus replication in vitro [17,18]. The DNA-binding domain of NFI shows no obvious sequence similarity to any of the known classes of DNA-binding domains such as zinc finger, leucine zipper or helix–loop–helix motifs. Recently, a putative α-helical segment presumed mediating DNA contact was identified [19].

It has been demonstrated that NFI proteins derived from the human NFI-C/CTF gene, the rat NF1-L gene, and the porcine NFI-C/CTF gene bind to DNA as homodimers [6,17,18]. The question arises whether NFI proteins encoded by different NFI genes can form heterodimeric complexes. In this study, we have examined the heterodimerization and DNA binding potential of NFI dimeric transcription factor proteins encoded by the four chicken NFI genes.

2. Materials and methods

2.1. Construction of expression vectors

NFI cDNA inserts [7,8] were cloned into a Bluescript derived vector, pRB0.5, containing two copies of the 5′ nontranslated region of the rabbit β-globin gene [20]. The full-length NFI-A1 cDNA insert was excised with *Eco*RI, filled in with DNA polymerase I (Klenow) and ligated into pRB0.5 cleaved with *Xho*I and *Hin*dIII after Klenow fill in. The cNFI-B2 cDNA insert was transferred as an *Xho*I–*Hin*dIII fragment into the in vitro transcription vector. The cNFI-C2 insert was cloned as an *Xho*I–*Bam*HI fragment after partial *Bam*HI digestion of the parental vector. The cNFI-X construct lacking the first 9 amino acids was constructed in two steps. First, the construct encoding the truncated NFI-X protein, NFI-X(10–285), was prepared (see below). In the second step the 0.75 kb *Bsp*MI–*Xba*I fragment of this plasmid was replaced by the 1.3 kb *Bsp*MI–*Xba*I fragment of the original NFI-X cDNA clone p214.1. Vectors encoding truncated NFI polypeptides comprising the DNA binding and dimerization domain were amplified with suitable PCR-primers as described previously [7] and cloned into

*Corresponding author. *Present address:* The Scripps Research Institute, Department of Molecular and Experimental Medicine SBR-7, Division of Oncovirology, 10666 N. Torrey Pines Road, La Jolla, CA 92037, USA. Fax: (1) (619) 554-6919.

Abbreviations: NFI, nuclear factor I; CTF, CCAAT binding transcription factor; PCR, polymerase chain reaction; ORF, open reading frame.

SSDI 0014-5793(94)00585-J

the pRB0.5 vector. After cloning of the amplified fragments, NFI ORFs were verified by DNA sequencing. Oligonucleotides were synthesized by the solid phase phosphor-amidite method. Alternative nucleotides in brackets are present simultaneously.

3′ PCR-primer ON80 (amino acid 10 to 15 of cNFI-A1 and cNFI-X, position 11 to 16 of cNFI-C2 and cNFI-B2):
5′-GACTCGAGGCCGCCACCATGGATGA(A/G)TT(T/C)CA(T/C)CC(C/G)TT-3′.
Sequences of 3′ PCR-primers:
ON53 (cNFI-A1 amino acid 266 to 271):
5′-GGAATTCAAGAGGATGTGCTAGGTAA-3′;
ON52 (cNFI-B2 amino acid 263 to 268):
5′-GTAAGCTTTGTCAGCTTGGTGGAGAGGATAA-3′;
ON54 (cNFI-C2 amino acid 272 to 277):
5′-GGAATTCAGCTGGAGGAGGTGCTGGG-3′;
ON118 (cNFI-X amino acid 279 to 285):
5′-GGAATTCTCAGTCCCCCTCGTCCAGCGCCTT-3′.

2.2. In vitro transcription and translation

Linearized vectors were transcribed with T3 RNA polymerase in the presence of a CAP analog, $m^7G(5')ppp(5')G$, according to the protocol of the supplier (Stratagene). Nuclease-treated rabbit reticulocyte lysate (Promega) was programmed with 1 μl of RNA solution in the presence of [^{35}S]methionine (Amersham; 1000 Ci/mmol). Reactions were incubated for 90 min at 30°C and stored at −20°C. Aliquots of 2.5 μl were analysed by denaturing polyacrylamide gel electrophoresis [21]. Gels were dried and autoradiographed on Kodak XAR5-film.

2.3. Electrophoretic mobility shift assay

The probe consisted of a DNA fragment bearing the wild type (T2/WT) or a mutant (T2/PM) NFI binding site of the T2 region of the −6,1 kb chicken lysozyme gene enhancer [14]. In mutation T2/PM (PM = peripheral mutant), the wildtype T2 sequence was altered by replacing two G–C basepairs with T–A basepairs, thus eliminating two essential G(N)7 contacts for NFI: T2/WT, 5′-CTGGCACTATGCCAC-3′; T2/PM, 5′-CTTGCACTATGCAAC-3′. The mobility shift assay was performed as described previously [22]. 10 fmol of ^{32}P-endlabelled DNA probe was incubated with 1 μl of reticulocyte translation reaction programmed with NFI RNAs. The X-ray film was shielded of the ^{35}S-labelled proteins by two sheets of paper.

3. Results and discussion

3.1. Expression of full-length and truncated NFI proteins

NFI proteins bind as dimers to palindromic or half site DNA target sequences. To demonstrate that NFI polypeptides encoded by different NFI genes associate as heterodimers, we needed full-length and truncated NFI proteins. Inserts of full-length cDNA clones of cNF-A1, cNFI-B2 and cNFI-C2 [7] were cloned into an in vitro transcription vector (Fig. 1). This vector, pRB0.5, contains two copies of the 5′ nontranslated region of the rabbit β-globin gene allowing efficient in vitro translation. However, a construct containing the entire cDNA sequence of the cNFI-X type translated poorly, possibly due to the GC rich 5′ nontranslated sequence of the original cNFI-X clone (data not shown). Therefore the leader sequence including the first 9 codons was removed. For the synthesis of truncated NFI polypeptides, cDNA sequences comprising the DNA binding and dimerization domain were PCR-amplified and cloned into the in vitro transcription vector pRB0.5 (Fig. 1). In vitro translation of various NFI proteins yielded polypeptides of the expected size (Fig. 2). The rabbit reticu-

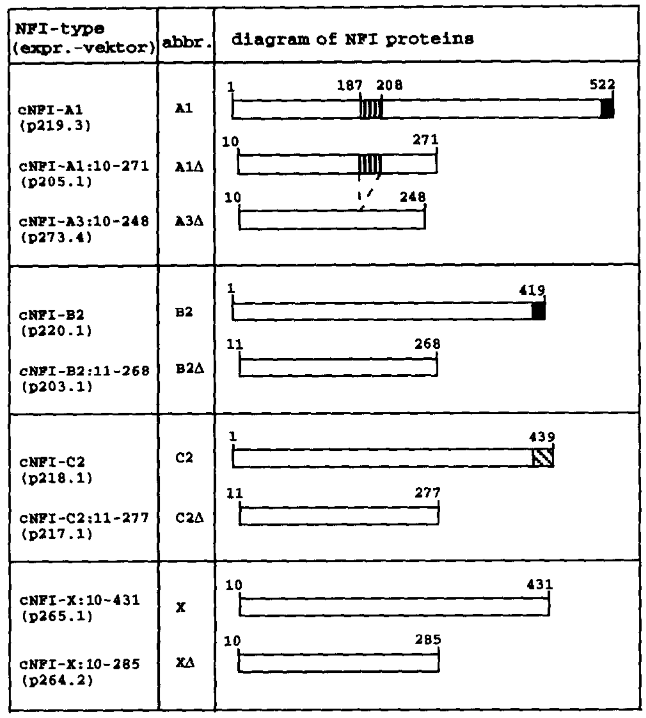

Fig. 1. Schematic representation of NFI gene constructs used for in vitro transcription/translation experiments. Within each group open boxes represent identical sequences. Domain 3, which is deleted in the NFI-A3 variant, is marked by vertical stripes. Carboxy terminal sequences characteristic for NFI isoforms are marked by black and striped boxes [24].

locyte in vitro translation system was used for synthesis of NFI proteins because it allows easy cotranslation of different RNA templates.

3.2. Homo- and heterodimerization of NFI proteins

In order to demonstrate the formation of NFI dimers, we followed the strategy described by Hope and Struhl [23]. Comparable amounts of full-length cNFI-A1 and truncated cNFI-A1, cNFI-B2, cNFI-C2 and cNFI-X polypeptides were translated separately or cotranslated, and shown to bind a DNA fragment comprising the T2 region of the −6.1 kb enhancer of the chicken lysozyme gene [14]. This sequence harbours a palindromic DNA binding sequence closely resembling the NFI consensus binding site 5′-YTGGCA(N)$_3$TGCCAR-3′, with the exception of the 3′ flanking purine which is substituted by an pyrimidine. As a control for sequence specific DNA binding we used a mutant binding site, T2/PM, in which at least two essential G(N)7 H-bonds between DNA and the protein were eliminated [14].

Analysis of the unprogrammed reticulocyte lysate detected a high molecular weight NFI complex endogenous to the system which was competed by in vitro translated proteins (Fig. 3, lanes 1 and 3). The lysate programmed with trancripts encoding full-length cNFI-C2 proteins gave rise to a protein–DNA complex in the upper part

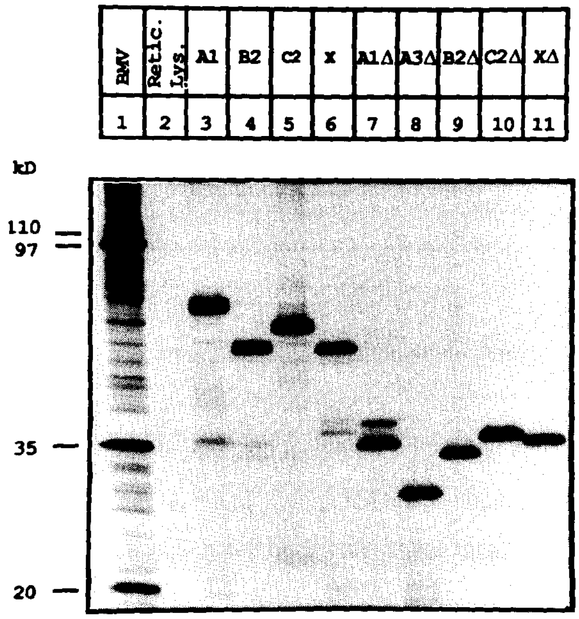

Fig. 2. Autoradiograph of polyacrylamide gel electrophoresis of in vitro translated ^{35}S-labelled NFI proteins. Lane 1, brome mosaic virus proteins serving as molecular weight markers; lane 2, unprogrammed reticulocyte lysate. For abbreviations see Fig. 1.

of the gel, whereas the truncated cNFI-C2 protein yielded a more rapidly migrating complex (Fig. 3, lanes 3 and 17). Cotranslation of long and short cNFI-C2 proteins resulted in an additional band (Fig. 3, lane 21). This intermediate band shows that polypeptides of different length can associate, indicating that NFI-C2 proteins can form homodimers. However, if long and short proteins were translated separately and mixed afterwards, no band of intermediate size was detectable (Fig. 3, lane 19). Similarly, cotranslation of full-length cNFI-C2 and truncated cNFI-A1, cNFI-B2 and cNFI-X proteins gave rise to intermediate bands in the mobility shift gel (Fig. 3, lanes 9, 15 and 27). This result demonstrates that proteins derived from different NFI genes can form heterodimers. Sequential translation followed by mixing of lysates does not result in heterodimer formation (Fig. 3, lanes 7, 13 and 25). We combined systematically each full-length NFI protein with the four truncated proteins according to the scheme shown in Fig. 3 (NFI-B2, Fig. 4 A; NFI-X, Fig. 4 B; NFI-A1, Fig. 4 C). The full-length cNFI-A1 protein produced a broad band (Fig. 4 C, lane 3), probably due to partial proteolysis during the incubation of the DNA binding reaction. The DNA binding assay reveals that all combinations resulted in formation of heterodimers, when proteins were cotranslated but not when successively translated. Moreover, the reverse combinations of short cNFI-C2 protein with long cNFI-A1, cNFI-B2 and cNFI-X proteins crossvalidate the results shown in Fig. 3.

These experiments clearly show that the exchange of protein subunits in dimeric NFI complexes is extremely slow. It was demonstrated that the human NFI-C/CTF and the rat NF1-L protein can be crosslinked in solution in the absence of DNA [17,18]. In addition, the rat NF1-L protein binds as a dimer even to DNA half-sites, ruling

Protein	Retic. Lys.		C2		A1Δ		C2 A1Δ mix.		C2 A1Δ cotr.		B2Δ		C2 B2Δ mix.		C2 B2Δ cotr.		C2Δ		C2 C2Δ mix.		C2 C2Δ cotr.		XΔ		C2 XΔ mix.		C2 XΔ cotr.	
DNA: T2	WT	PM	WT	PM	WT	PM	WT	PM	WT	PM	WT	PM	WT	PM	WT	PM	WT	PM	WT	PM	WT	PM	WT	PM	WT	PM	WT	PM
Lane	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28

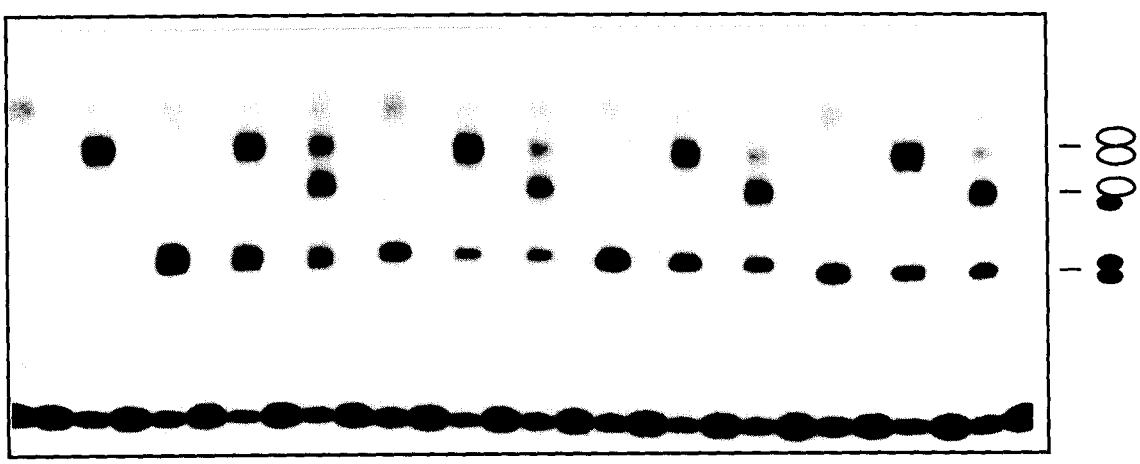

Fig. 3. Homo- and heterodimerization of NFI-C2 protein. Elecrophoretic mobility shift assay of DNA binding by full-length and truncated NFI polypeptides. DNA binding reactions contained unprogrammed lysate (lanes 1, 2), full-length C2 protein (lanes 3, 4), truncated C2 protein (lanes 17, 18), a mixture of both proteins following separate translations (mix., lanes 19, 20), or full-length C2 protein cotranslated with truncated C2 protein (cotr., lanes 21, 22). The other combinations of full-length C2 and truncated A1, B2 and X proteins were tested accordingly. The DNA probe comprised the T2 element of the −6.1 kb enhancer of the chicken lysozyme gene (WT, wildtype sequence; PM, peripheral mutant).

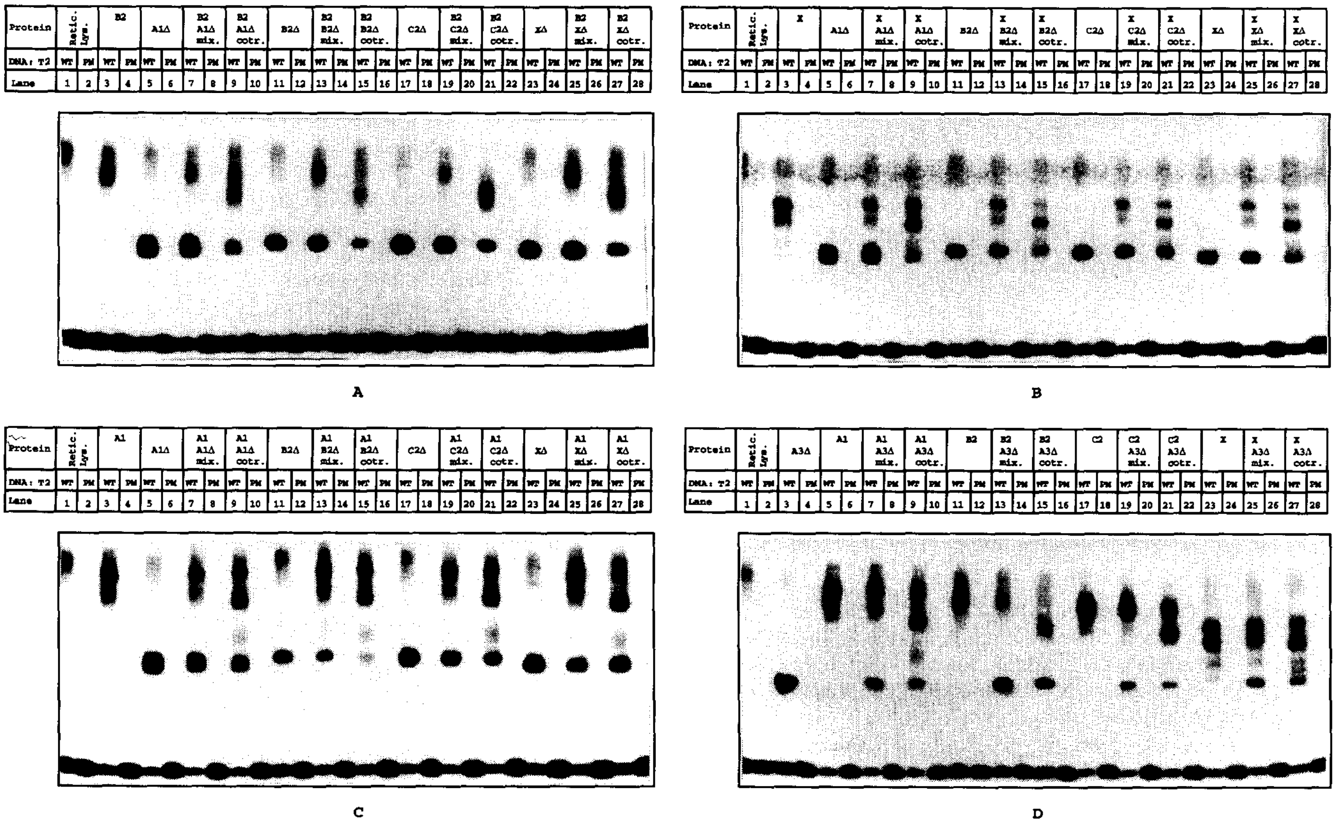

Fig. 4. Elecrophoretic mobility shift assay of DNA binding by combinations of full-length and truncated NFI polypeptides. Homo- and heterodimerization of (A) NFI-B2, (B) NFI-X, (C) NFI-A1 and (D) NFI-A3 proteins. For experimental details see Fig. 3.

out the possibility that dimers can only assemble on palindromic target sequences. Although we were not able to detect exchange of NFI subunits in vitro, there might be cellular mechanisms inhibiting spontaneous assembly and/or facilitating disassembly and reassembly of NFI complexes in vivo.

3.3. Dimerization properties of the NFI-A3 isoform

To analyse the dimerization potential of NFI variants, we tested the recently isolated cNFI-A3 isoform which lacks 22 amino acid residues within the DNA binding domain as compared to cNFI-A1 (amino acid residues 187–208 of cNFI-A1 [24]). Similar to the truncated cNFI-A1 protein, the cNFI-A3 variant formed heterodimeric complexes with cNFI-A1, cNFI-B2, cNFI-C2 and cNFI-X (Fig. 4D). Again, the formation of mixed dimers was dependent on cotranslation (Fig. 4D, lanes 9, 15, 21 and 27). In contrast, no additional bands were detected when the proteins were synthesized sequentially (Fig. 4D, lanes 7, 13, 19 and 25). This result indicates that the naturally occurring 22 amino acid residue deletion does not significantly affect dimerization or DNA binding. However, the protein domain N-terminal to position 186, which is identical in both variants, is by itself not sufficient for DNA binding [18]. Amino acids between position 209 and 271 of the cNFI-A1 protein are necessary for DNA binding. This result implies that in the cNFI-A1 protein noncontiguous sequences are necessary for DNA binding.

The 22 amino acid stretch missing in the cNFI-A3 isoform is mainly composed of hydrophilic amino acid residues, displaying a net negative charge. We speculate that this sequence is not an integral part of the DNA binding domain, but located on the surface of the molecule, providing an interface for the interaction with other proteins or for posttranslational modification.

3.4. The diversity of the NFI family

The combinatorial association of NFI gene products reveals a large variety of possible dimeric NFI complexes, further increasing the diversity of the NFI protein family. In total, we have isolated cDNA sequences of six NFI-A, two NFI-B, three NFI-C isoforms, and one NFI-X isoform, adding up to a total of 12 different chicken NFI gene products [24]. In this study, we included only one representative isoform from each NFI gene. Since the non-tested protein isoforms derived from each gene share identical DNA-binding and dimerization domains as the tested ones, it is most likely that they too would homo- and heterodimerize. At present, we cannot rule out that truncation of one subunit is required for unrestricted heterodimerization. We have no direct evidence for heterodimerization in intact cells, but Northern blot analysis of RNA from chicken cell lines shows that various combinations of NFI genes are simultaneously expressed, which provides preconditions necessary for heterodimerization (F. Qian, U. Kruse and A.E. Sippel, unpublished results). It remains to be determined which combinations of NFI homo- and heterodimers are present in a given cell type and whether they are functionally distinct in terms of their transcriptional activation properties.

Acknowledgements: This work was supported by grants to A.E.S. from the Deutsche Forschungsgemeinschaft (Si165/7–1), the Fonds der Chemischen Industrie, and the German Israeli Foundation for Scientific Research and Development (GIF). U.K. was a fellow of the Studienstiftung des deutschen Volkes. We thank Gabor Igloi for the synthesis of oligonucleotides. U.K. thanks Peter K. Vogt and Eric R. Schuur for critical reading of the manuscript.

References

[1] Struhl, K. (1991) Neuron 7, 177–181.
[2] Hirai, S.-I., Ryseck, R.-P., Mechta, F., Bravo, R. and Yaniv, M. (1989) EMBO J. 8, 1433–1439.
[3] Santoro, C., Mermod, N., Andrews, P.C. and Tjian, R. (1988) Nature 334, 218–224.
[4] Paonessa, G., Gounari, F., Frank, R. and Cortese,R. (1988) EMBO J. 7, 3115–3123.
[5] Gil, G., Smith, J.R., Goldstein, J.L., Slaughter, C.A., Orth, K., Brown, M.S. and Osborne, T.F. (1988) Proc. Nat. Acad. Sci. USA 85, 8963–8967.
[6] Meisterernst, M., Rogge, L., Foeckler, R., Karaghiosoff, M. and Winnacker, E.-L. (1989) Biochem. 28, 8191–8200.
[7] Rupp, R.A.W., Kruse, U., Multhaup, G., Göbel, U., Beyreuther, K. and Sippel, A.E. (1990) Nucleic Acids Res. 18, 2607–2616.
[8] Kruse, U., Qian, F. and Sippel, A.E. (1991) Nucleic Acids Res. 19, 6641.
[9] Jackson, S.P. and Tjian, R. (1988) Cell 55, 125–133.
[10] Lamb, P. and McKnight, S.L. (1991) Trends Biochem. Sci. 16, 417–422.
[11] Borgmeyer, U., Nowock, J. and Sippel, A.E. (1984) Nucleic Acids Res. 12, 4295–4311.
[12] Martinez, E., Dusserre, Y., Wahli, W. and Mermod, N. (1991) Mol. Cell. Biol. 11, 2937–2945.
[13] Graves, R.A., Tontonoz, P., Ross, S.R. and Spiegelmann, B.M. (1991) Genes Dev. 5, 428–437.
[14] Grewal, T., Theisen, M., Borgmeyer, U., Grussenmeyer, T., Rupp, R.A.W., Stief, A., Qian, F., Hecht, A. and Sippel, A.E. (1992) Mol. Cell. Biol. 12, 2339–2350.
[15] Nagata, K., Guggenheimer, R.A., Enomoto, T., Lichy, J. H. and Hurwitz, J. (1982) Proc. Nat. Acad. Sci. USA 79, 6438–6442.
[16] Leegwater, P.A.J., van der Vliet, P.C., Rupp, R.A.W., Nowock, J. and Sippel, A.E. (1986) EMBO J. 5, 381–386.
[17] Mermod, N., O'Neill, E.A., Kelly, T.J. and Tjian, R. (1989) Cell 58, 741–753.
[18] Gounari, F., De Francesco, R., Schmitt, J., van der Vliet, P.C., Cortese, R. and Stunnenberg, H. (1990) EMBO J. 9, 559–566.
[19] Suzuki, M. (1993) EMBO J. 12, 3221–3226.
[20] Annweiler, A., Hipskind, R.A. and Wirth, T. (1991) Nucleic Acids Res. 19, 3750.
[21] Laemmli, U.K. (1970) Nature 227, 680–685.
[22] Rupp, R.A.W. and Sippel, A.E. (1987) Nucleic Acids Res. 15, 9707–9726.
[23] Hope, I.A. and Struhl, K. (1987) EMBO J. 6, 2781–2784.
[24] Kruse, U. and Sippel, A.E. (1994) J. Mol. Biol. 238, 860–865.