The EMBO Journal vol.7 no.10 pp.3115-3123, 1988

Purification of a NF1-like DNA-binding protein from rat liver and cloning of the corresponding cDNA

Giacomo Paonessa, Fotini Gounari, Rainer Frank and Riccardo Cortese

European Molecular Biology Laboratory, Meyerhofstrasse 1, 6900 Heidelberg, FRG

Communicated by R.Cortese

NF1-like proteins play a role in transcription of liver-specific genes. A DNA-binding protein, recognizing half of the canonical NF1 binding site (TGGCA) present on the human albumin and retinol-binding protein genes, has been purified from rat liver. Several peptides deriving from a tryptic digest of the purified protein were sequenced and the sequence was used to synthesize specific oligonucleotides. Two overlapping cDNA clones were obtained from a rat-liver cDNA library; their sequence reveals an open reading frame coding for 505 amino acids, including all the peptides sequenced from the purified protein. The DNA-binding domain, most likely located within the first 250 amino acids, is highly homologous to the sequence of CTF/NF1 purified from HeLa cells. Northern analysis reveals several mRNA species present in different combinations in various rat tissues.

Key words: albumin/NF1-L cDNA/rat liver/retinol-binding protein

Introduction

In recent years there has been remarkable progress in our understanding of the basic mechanisms responsible for the regulation of cell-type-specific gene expression (for a review see Maniatis *et al.*, 1987). A major breakthrough occurred when it was shown that cloned segments of DNA introduced into cultured cells displayed the cell-type-specific expression observed *in vivo* (Banerji *et al.*, 1983; Gillies *et al.*, 1983; Neuberger, 1983; Queen and Baltimore, 1983; Walker *et al.*, 1983). These results indicate that differentiated cells contain a specialized transcriptional apparatus; in most cases, the information required for specificity of expression is contained within relatively short DNA segments. In many systems it has been possible to define the cell-specific *cis*-acting elements and, with a combination of DNA-binding assays and *in vitro* transcription experiments, to identify and purify cell-specific *trans*-acting factors (reviewed in Jones *et al.*, 1988). How is a specialized transcriptional apparatus established and maintained in differentiated cells? It is most likely that, in order to answer this question, it will be necessary to clone the genes coding for the cell-specific *trans*-acting factors. The clones will be invaluable new reagents with which to investigate the nature and the activity of the specialized transcriptional apparatus during development.

Several liver-specific genes have been cloned and their expression studied by us and others in cultured cell lines (Ott *et al.*, 1984; Ciliberto *et al.*, 1985; D'Onofrio *et al.*, 1985; Schaul *et al.*, 1985; Godbout *et al.*, 1986; Muglia and Rothman-Denes, 1986; Babiss *et al.*, 1987; Colantuoni *et al.*, 1987; Courtois *et al.*, 1987; De Simone *et al.*, 1987; Heard *et al.*, 1987; Lee *et al.*, 1987; Oliviero *et al.*, 1987; Shelley and Baralle, 1987; Godbout *et al.*, 1988; Grayson *et al.*, 1988a,b), in transgenic mice (Ciliberto *et al.*, 1987; Kelsey *et al.*, 1987; Pinkert *et al.*, 1987; Ruther *et al.*, 1987; Dente *et al.*, 1988) and *in vitro* (Gorski *et al.*, 1987; Lichtsteiner *et al.*, 1987; Hardon *et al.*, 1988; Kugler *et al.*, 1988; Monaci *et al.*, 1988). This work has led to the characterization of several DNA-binding proteins present in liver nuclei which, in different combinations, bind to different liver-specific genes. Some *trans*-acting factors present in the liver are absent from other tissues (Courtois *et al.*, 1987; Lichtsteiner *et al.*, 1987; Grayson *et al.*, 1988b; Monaci *et al.*, 1988), while others appear to be ubiquitous (Cereghini *et al.*, 1987; Lichtsteiner *et al.*, 1987; Monaci *et al.*, 1988). Among these factors a DNA-binding activity is present which recognizes the sequence TGGCA, previously shown to be the recognition site of the DNA-binding protein NF1 and present on several liver-specific genes (Borgmeyer *et al.*, 1984; Shaul *et al.*, 1986; Lichtsteiner *et al.*, 1987; Raymondjean *et al.*, 1988). Analysis of the protein(s) interacting with this sequence in rat-liver nuclear extracts has shown that the situation is rather complex as different NF1-like species (forms), distinguishable by mobility retardation experiments, are present only in liver while others are also present in other tissues (Lichtsteiner *et al.*, 1987). As a first step in establishing the role of NF1-like activities in liver-specific gene expression we have purified a TGCCA-binding protein from rat liver, referred to as NF1-L(iver), and have cloned the corresponding cDNA. Using the cDNA as a probe, we observe a considerable degree of heterogeneity in NF1 transcripts and a characteristic tissue-specific expression pattern.

Results

Purification of a TGGCA-binding protein from rat liver

A DNA-binding activity present in rat liver nuclear extracts, whose importance for transcription has been previously established (Ott *et al.*, 1984; Gorski *et al.*, 1986; Colantuoni *et al.*, 1987) interacts with the 5′ flanking regions of the retinol-binding protein (*Rbp*) and albumin (*Alb*) genes. We fractionated a rat-liver nuclear extract by chromatography on Heparin–Sepharose and MonoS Fast Flow columns. The sequence specificity of the purified DNA-binding factor was assayed by DNase I protection experiments on the human *Rbp* and the *Alb* promoters as probes (Figure 1). The footprint on the *Rbp* gene is located between N-271

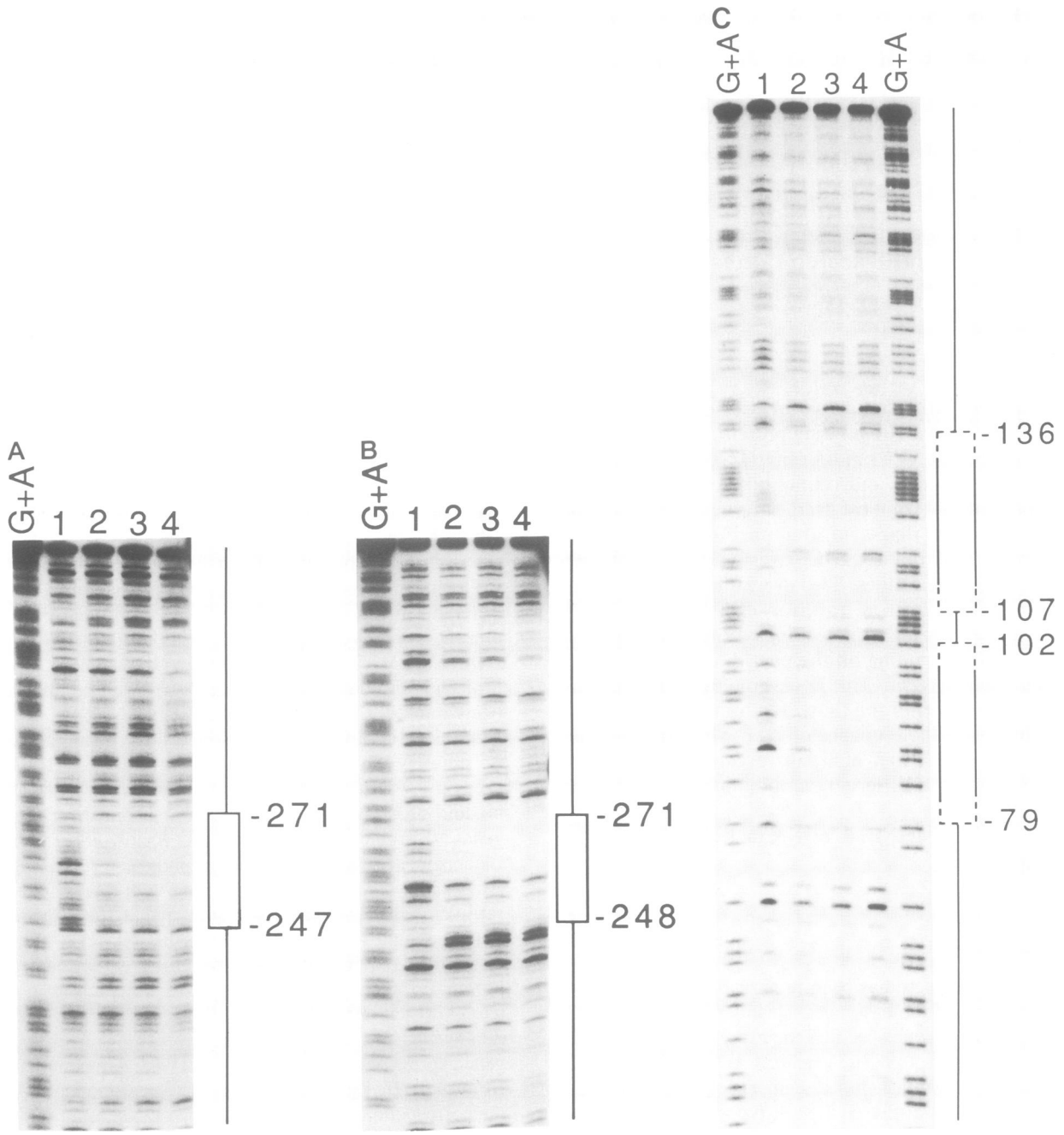

Fig. 1. DNase I footprints of affinity-purified TGGCA-binding protein on the *Rbp* and *Alb* genes. G + A chemical cleavages of the end-labelled probes were used as sequence markers. Boxes alongside the autoradiographs indicate the regions protected from DNase I digestion; numbers refer to the positions relative to the start site of transcription. (**A**) Footprint on the bottom strand of the *Rbp* promoter region. (**B**) Footprint on the top strand of the *Rbp* promoter region. (**C**) Footprint on the top strand of the *Alb* promoter region. In each panel: **lane 1**, digestion of naked DNA; **lanes 2–4**, digestion in the presence of 4, 10, 20 ng of purified protein.

and N-248 (Figure 1, panels A and B). In the case of the *Alb* promoter there are two footprints, one between N-136 and N-107, and a second, observed only at higher protein concentrations, between N-102 and N-79 (Figure 1, panel C). In both cases the boundaries of the footprints are not well defined due to the low frequency of DNase I cutting in these regions. Synthetic oligonucleotides were synthesized on the basis of the footprint data (Figure 2, panel A). alb1, spanning the region from N-137 to N-76 of the *Alb* promoter, yields two shifted bands in mobility retardation assays. alb2 and alb3, which separately span the two adjacent footprints, show a single, shifted band. The two shifted bands obtained with alb1 might represent protein–DNA complexes on one or two sites respectively. The oligonucleotide rbp yields a shifted band with the same mobility as the alb2 and alb3 complexes (Figure 2, panel B). Cross competitions in mobility retardation assays and determination of the contact points by methylation interference experiments suggest that the same protein is responsible for the formation of the complex with the four oligonucleotides (data not shown). The three identified binding sites share the sequence TGGCA.

The protein purified from rat liver interacts with DNA in a manner very similar to the well-characterized CTF/NF1

Fig. 2. Mobility retardation assay of the purified protein with different oligonucleotides. **Panel A** shows the *Alb* and *Rbp* double-stranded synthetic oligonucleotides used. Asterisks indicate the contact points as determined by methylation interference experiments (data not shown). **Panel B** shows the result of the gel-retardation experiment using as probe: alb1, in **lanes 1** and **2**; alb2 in **lanes 3** and **4**; alb3 in **lanes 4** and **5**; rbp in **lanes 6** and **7**. Plus and minus indicate the absence or presence of the purified protein.

protein purified from HeLa cells (Jones *et al.*, 1987). NF1 purified from HeLa cells binds to the mouse and rat *Alb* promoter in a region that is highly homologous to the region −137/−107 of the human *Alb* promoter region (for a direct comparison, see Cereghini *et al.*, 1987). In the human *Alb* promoter we observe a second binding site at position N-102/-79.

It was possible to purify this TGGCA binding activity further by DNA-affinity chromatography (see Materials and methods). A SDS−PAGE of the last step of the purification is shown in Figure 3, panel A. The most abundant species is a protein of ~30 kd. We have established that the DNA-binding activity co-migrates with this protein by a denaturation−renaturation experiment (Figure 3, panel B). Several slices were cut from the SDS−PAGE, the proteins were eluted and renatured according to the protocol described in Materials and methods and tested in a mobility retardation assay, using the alb1 oligo as a probe. The DNA-binding activity can be recovered from the gel slices containing the 30 000 mol. wt protein. In addition the activity is associated with most of the molecular species present in the 30-kd band, as shown by a preparative gel-shift experiment. We first fractionated the protein−DNA complex on native PAGE, together with control samples without DNA or without proteins. The region of the gel containing the specific protein−DNA complex, as well as the corresponding regions of the control lanes, was excised and applied on an SDS−PAGE. The results are shown in Figure 4, panel A. Lane 1 contains the protein sample used for the assay; lanes 2, 4 and 5 the controls without DNA, with unrelated DNA and without protein respectively; lane 3 contains the sample derived from the shifted band. It appears that the major specie present in our most purified preparation, a 30 000 mol. wt polypeptide, is the DNA-binding activity. We refer to it as NF1-L.

Microsequencing of NF1-L and preparation of specific antibodies

The protein was eluted from a preparative SDS−PA gel and

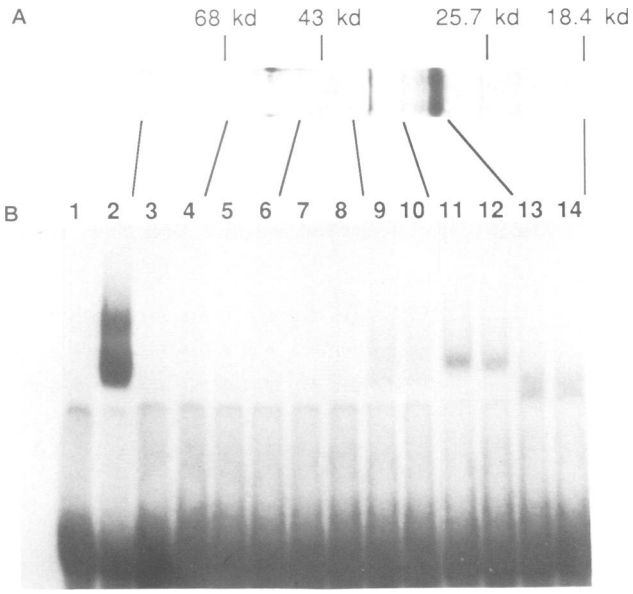

Fig.3. Renaturation of the TGCCA binding protein. Affinity-purified TGGCA-binding protein was fractionated on an SDS−PAGE. Six slices were cut from the gel and treated as described in Materials and methods. Samples were assayed by mobility retardation, using oligonucleotide alb1 as probe. (**A**) Silver stain of the purified protein (the indicated sections represent slices cut from a parallel lane). (**B**) Mobility retardation assay with the renatured proteins. **Lane 1**, alb1; **lane 2**, alb1 in the presence of the affinity-purified material; in the following odd lanes the mobility retardation mixtures with material from slices 1−6 was applied in the absence of salmon sperm DNA; in the even lanes are the same mixtures in the presence of 100 ng of salmon sperm DNA.

used for microsequencing. Tryptic digests yielded several peptides, five of which were purified and sequenced (see Materials and methods). Their amino acid sequence is reported in Figure 5, panel A. We chemically synthesized the oligopeptides corresponding to the sequences 1, 3 and 6, coupled them to a carrier protein and injected them into different rabbits. Peptide 6 turned out to be strongly

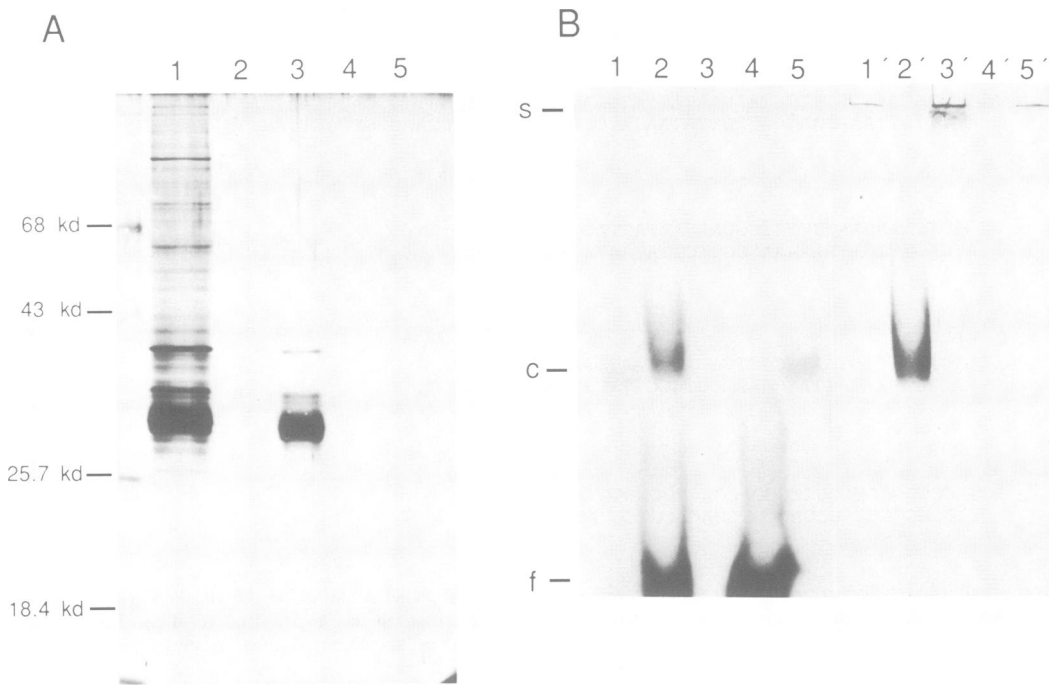

Fig. 4. Preparative mobility retardation experiments. (**A**) The MonoS fraction containing the TGGCA-binding protein activity was fractionated on native PAGE, in the presence or absence of alb2 oligo. The shifted band was visualized by autoradiography and run on an SDS−PAGE. The corresponding area of the gel from the slot with protein but without DNA was also processed as a control. In the figure we show the results of the SDS−PAGE, with proteins revealed by silver staining. **Lane 1**, active MonoS fraction; **lane 2**, 2nd dimension of control protein without DNA; **lane 3**, 2nd dimension of protein; **lane 4**, 2nd dimension of protein + unrelated DNA; **lane 5**, 2nd dimension of the DNA without protein. (**B**) Preparative mobility retardation and immunostaining of the shifted protein. Gel-retardation mixtures with or without DNA probe were fractionated on a native PAGE. The gel was electroblotted onto a membrane to retain the proteins, followed by DEAE−cellulose, to retain the DNA. The DNA was revealed by autoradiography; **lanes 1** and **5**, mobility retardations with analytical amounts of protein and labelled alb2; **lane 2**, shift obtained in the presence of preparative amounts of DNA and protein; **lane 3**, proteins in the absence of DNA; **lane 4**, DNA in the absence of proteins. Proteins were revealed by immunostaining with anti-pep-6: **lanes 1′−5′**, as above.

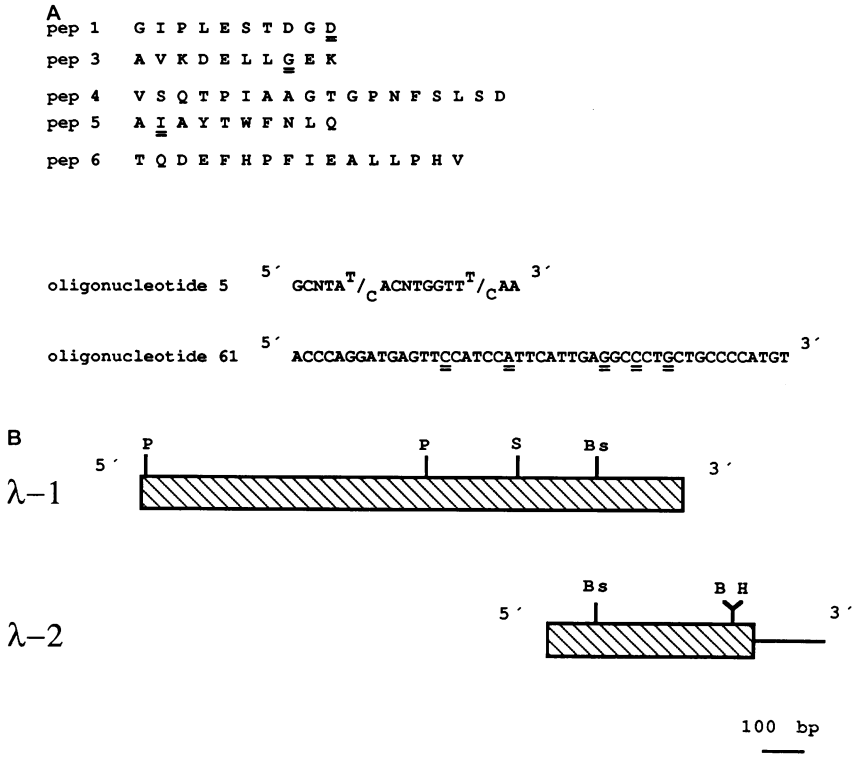

Fig. 5. (**A**) Amino acid sequence of the five peptides derived from tryptic digestion of the purified protein. Oligonucleotides 5 and 61 are the single-stranded probes used in the screening of the library. Amino acids or nucleotides that turned out to be different in the cDNA are double underlined. (**B**) Schematic representation of the clones λ-1 and λ-2. P, *Pst*I; S, *Sph*I; Bs, *Bst*xI; B, *Bam*HI; H, *Hin*dIII.

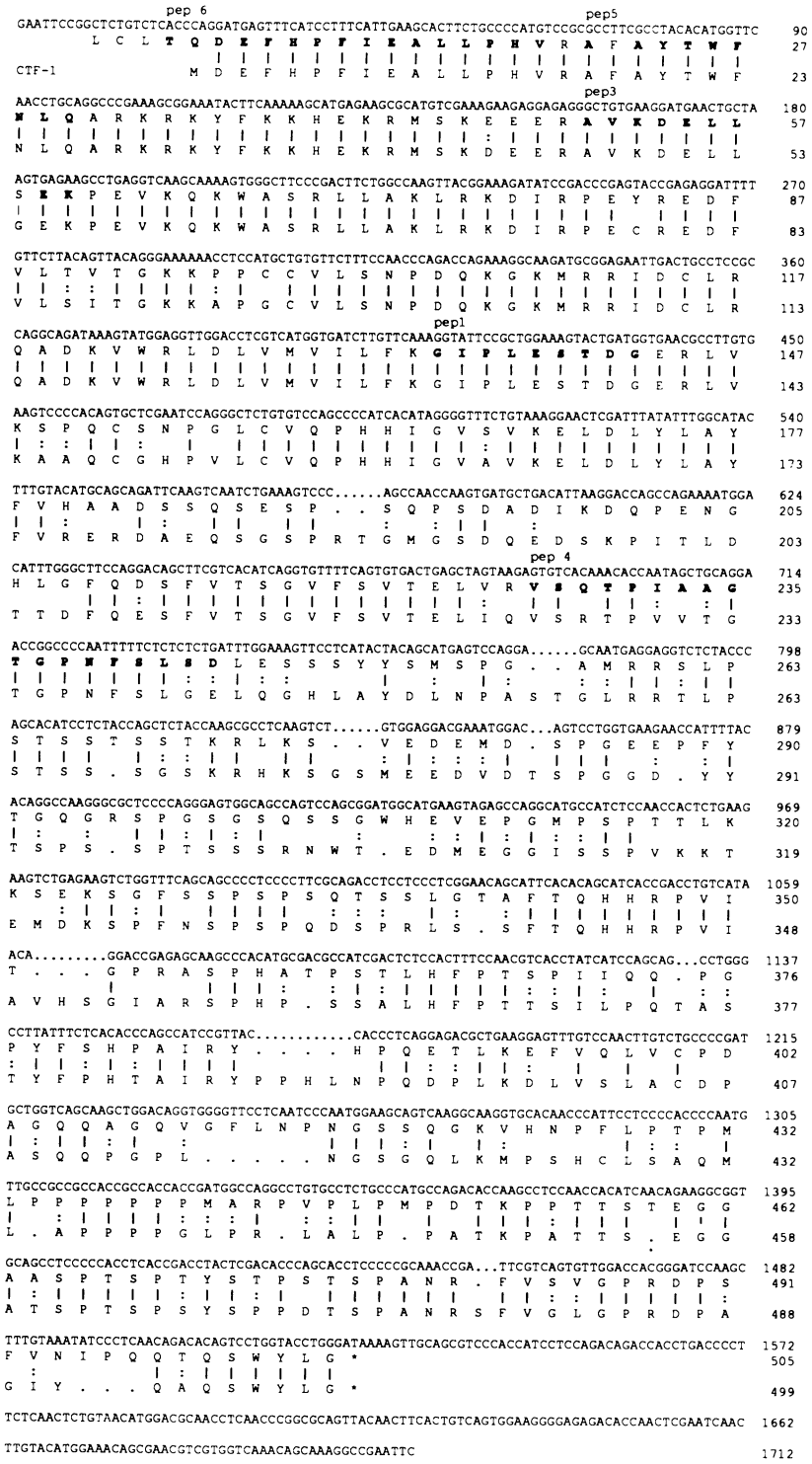

Fig. 6. Sequence of NF1-L and comparison with CTF-1 at protein level. Gaps introduced to maximize the homology are represented by dots. Vertical bars highlight identical residues, double dots indicate conservative amino acid substitution, grouped as follows: (A, P, G, S, T), (Q, N, E, D), (V, I, M, L), (K, R, H), (F, Y, W) (Schwartz and Dayhoff, 1978).

antigenic. The corresponding antibody reacts, in a Western blot, with a 30-kd protein, comigrating with the species responsible for the DNA-binding activity (data not shown). In order to prove that the antibody reacts with the DNA-binding protein we performed a preparative gel-shift experiment, separating proteins in the presence or absence of DNA probe, on a native gel. The gel was electroblotted onto an Immobilon membrane, which retains proteins, whereas the DNA was collected onto a sheet of DE 52 paper layered behind the membrane. The DNA on the DE 52 paper was visualized by autoradiography (Figure 4, panel B; lanes 1 and 5 contain analytical amounts of the reaction mix; lane 2 preparative amounts of the reaction mix, lanes 3 and 4 the negative controls without DNA and without proteins respectively). The proteins were visualized with specific serum. The results (Figure 4, panel B) show that the protein

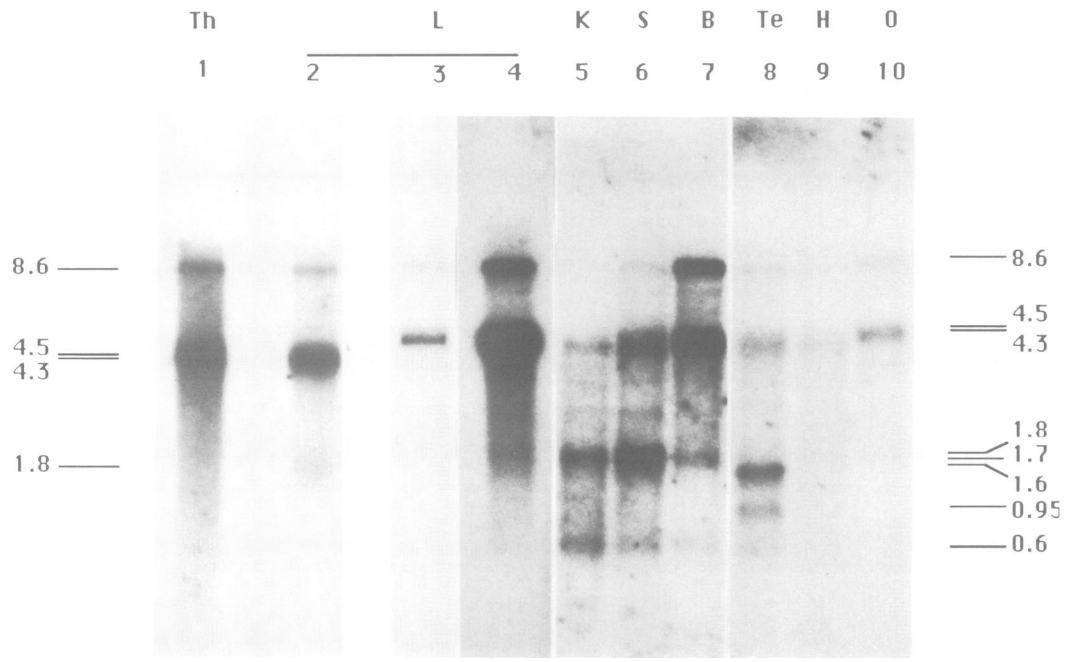

Fig. 7. Northern blot on poly(A)$^+$ RNA extracted from several rat tissues. **Lane 1**, thyroid; **lanes 2–4**, liver; **lane 5**, kidney; **lane 6**, spleen; **lane 7**, brain; **lane 8**, testis; **lane 9**, heart; **lane 10**, ovary.

recognized by the antibody does not enter the native gel in the absence of DNA (slot 3′ for the preparative assay, 1′ and 5′ for the analytical assay) but is shifted into the gel in a complex with specific DNA (slot 2′). Thus we conclude that the anti-pep6 serum reacts with NF1-L. Attempts to use the antibody to interfere with the DNA-binding activity failed, most likely because the antibody does not react with the native protein.

Cloning of the cDNA coding for NF1-L

The sequences of peptides 5 and 6 were used to synthesize oligonucleotides (Figure 5, panel A) to screen a rat liver cDNA library. Oligonucleotide 6 was synthesized as a single DNA segment, taking into account codon usage in mammals. Oligonucleotide 5 was synthesized as a mixture of all possible sequences. A rat liver λgt11 cDNA library (see Materials and methods) was screened first with oligonucleotide 6. Several positive clones were obtained, which were then screened with oligonucleotide 5. Three clones hybridized to both probes. All three contained an identical insert of 1350-bp DNA and were called λ-1. Using the DNA from λ-1 as probe we identified another clone, λ-2, whose sequence overlaps with that of λ-1 but extends further downstream. The combined sequence of both inserts contains an open reading frame of 505 amino acids. As there is no poly(A) tail and probably no methionine initiation codon, the cDNA is incomplete at both ends. A schematic view is shown in Figure 5, panel B. The DNA sequence obtained from λ-1 and λ-2 is shown in Figure 6, together with the comparison of related sequences obtained by Santoro *et al.* (1988). The significance of these homologies will be discussed.

The gene(s) coding for NF1-L yields multiple transcripts

Mobility retardation experiments reveal the existence of multiple complexes of DNA containing the TGGCA sequence and proteins present in the rat liver nuclear extracts, some of which are absent from spleen nuclear extracts (Lichtsteiner *et al.*, 1987; P.Monaci and A.Nicosia, unpublished results). It is therefore possible that there are several TGGCA-binding proteins. The results of a Northern analysis (Figure 7) support this hypothesis. Poly(A)$^+$ RNA from several rat tissues were hybridized with ^{32}P-labelled λ-1 cDNA insert. At least eight RNA species are present, in a characteristic combination and relative abundance in each tissue. Assuming that co-migrating species correspond to identical RNAs, we have tabulated the specific pattern in each tissue (Table I). In liver, the 4.3-kb RNA is the major species, in brain and thyroid the major species are the 4.5- and the 4.3-kb RNA. In the ovary the 4.5-kb RNA is the major species. In spleen and kidney the major species is the 1.7-kb RNA, whereas in testis the major species is the apparently testis-specific 1.6-kb RNA. Also in liver, the 1.8-kb is a minor but apparently liver-specific RNA. We cannot distinguish whether the multiple transcripts derive from different genes or from different transcriptional units of the same gene or whether they are produced by alternative processing of a common precursor RNA.

Table I. Distribution of the major NF1 mRNA transcripts in different organs

	kb							
	8.6	4.5	4.3	1.8	1.7	1.6	0.95	0.6
Thyroid	+	+	+	–	–	–	–	–
Liver	+	–	+	+	–	–	–	–
Kidney	–	–	+	–	+	–	–	+
Spleen	+	–	+	–	+	–	–	+
Brain	+	+	+	–	+	–	–	+
Testis	+	–	+	–	–	+	+	+
Heart	–	–	+	–	–	–	–	–
Ovary	+	+	–	–	–	–	–	–

Discussion

Sippel and co-workers originally identified and characterized a DNA-binding activity, present in several chicken tissues, which recognized the sequence TGGCANNNTGCCA (Borgmeyer *et al.*, 1984). Thereafter several groups have purified similar activities from various different cell lines and shown that they bind to several genes, either to the canonical sequence or to a half site (TGGCA). Such binding sites appear to play role both in transcription and in replication (for a review see Santoro *et al.*, 1988). Mobility retardation experiments using the purified protein indicate a considerable degree of heterogeneity in the complexes formed (Jones *et al.*, 1987; Rosenfeld and Kelly, 1986; Santoro *et al.*, 1988) suggesting that there might be several different proteins with similar DNA-binding properties. For the sake of simplicity we will refer to these as NF1 activities, adopting the terminology originally used by Nagata *et al.* (1983).

NF1-like activities, recognizing the half site TGGCA, have been detected in rat-liver nuclear extracts (Cereghini *et al.*, 1987; Lichtsteiner *et al.*, 1987; Monaci *et al.*, 1988). We have purified a protein of ~30 kd,, which we call NF1-L, which recognizes the same sites as NF1 purified from HeLa cells (Lichtsteiner *et al.*, 1987; Raymondjean *et al.*, 1988). In the light of our results with the cDNA clone and with the antibody (G.Paonessa, unpublished results) we believe that the 30-kd species is a cleavage product of larger polypeptides, still containing the DNA-binding domain. The cDNA sequence shows that pep-6 is not preceded by lysine or arginine, indicating that it is probably at the amino terminus of the 30-kd protein. The most carboxy-terminal one is pep-4, ending at residue 244 of our sequence (Figure 6). Together this accounts for ~27 kd, corresponding almost to the mol. wt of the purified protein, and indicates that the DNA-binding domain is contained within the first 250 amino acids.

Independently Santoro *et al.* (1988) have isolated and sequenced three distinct cDNA clones coding for NF1 (also referred to as CTF) proteins present in HeLa cells. The comparison of the sequence of the rat cDNA from liver and the cDNAs from HeLa cells shows some interesting differences. It appears that the rat-liver cDNA codes for a protein with a different sequence for the first available five amino acids followed by a stretch of 175 amino acids which are practically identical to the human protein, with a homology of identical residues of 98%. The remaining part of the open reading frame shows an overall homology of 72%. The proteins coded by the HeLa and the rat liver cDNAs are thus very conserved in the region corresponding to the DNA-binding domain, while they diverge in the remaining sequence.

Northern analysis reveals multiple transcripts which react with the cDNA (Figure 7). The patterns are qualitatively and quantitatively tissue specific (Table I). At the moment we do not know to what extent the various transcripts code for functionally related proteins.

Since we do not yet have any information about the genomic sequence, we cannot exclude the possibility that there are multiple genes coding for different forms of NF1. Another possibility, rather more plausible in the light of the results of Santoro *et al.* (1988), and because of the blocks of homology and divergence between the rat and HeLa cell cDNAs, is that the diversity is generated by differential and tissue-specific RNA processing. In this view, it is interesting that the most amino-terminal part of the liver protein is joined to the first common segment at a region where, at the DNA level, there is a canonical splice acceptor junction (TCTCACCCAG/GA). In the alternative splicing hypothesis, this site is used as acceptor in the HeLa cell clones, bringing the methionine postulated as the initiator by Santoro *et al.* (1988) in the correct reading frame. In contrast, in the liver cDNA, the same region is apparently used as an exon, because it codes for the first amino acids (T and Q) present in pep-6 (Figure 5, panel A).

It has been established that the NF1 protein(s) has multiple functions, most notably as an activator of adenovirus replication (Nagata *et al.*, 1983; Leegwater *et al.*, 1985), and as a transcriptional activator (Jones *et al.*, 1987). In our laboratory we have evidence that it might also function as a repressor of transcription (P.Monaci and A.Nicosia, unpublished observations). It is reasonable to expect that these various functions might be performed by different NF1 species, constructed by assembling different domains in specific combinations. We imagine that a common domain, defined by the highly conserved sequence from amino acid 1 to ~245, is responsible for the binding to the TGGCA sequence; the other domains might be involved in different interactions with cellular components leading to activation or repression of transcription, or to DNA replication. *Trans*-acting factors which share common domains have been described in other systems: the adenovirus E1A region contains a single ORF that encodes both positive and negative transcriptional factors, generated by alternative splicing (Lillie *et al.*, 1986). Lambert *et al.* (1987) have shown that the BPV-1 E2 *trans*-activator and a BPV-1 transcriptional repressor share a carboxyl-terminal domain. Multifunctional proteins involved in transcription and DNA replication have been characterized in yeast (Shore and Nasmyth, 1987; Buchman *et al.*, 1988). It appears that the generation of different but related *trans*-acting factors plays an important role in the regulation of gene expression. We believe that the various mRNAs hybridizing with the NF1-L cDNA code for related *trans*-acting factors. Their characteristic pattern in each tissue (Table I) supports the idea that regulation is achieved through different combinations of a common set of regulatory molecules.

Materials and methods

Preparation of nuclear extracts and protein purification

Nuclear extracts were prepared from fresh rat liver essentially as described by Dignam *et al.* (1983). The crude nuclear extract was dialysed extensively against buffer D (20 mM Hepes, pH 7.9, 10% glycerol, 0.2 mM EDTA, 0.5 mM DTT, 1 mM PMSF, 100 mM KCl), and centrifuged for 10 min at 12 000 r.p.m. in a Sorvall SS-34 rotor. The clear supernatant was applied onto a DEAE–Sepharose column, and the flow through loaded onto a Heparin–Sepharose (HepS) column. The HepS column was eluted with a linear gradient of KCl (100 mM–2 M) in buffer D′, that contains only 0.1 mM PMSF and is supplemented with 0.05% NP-40; buffer D′ was used in all the subsequent steps. Fractions containing DNA-binding activities were pooled, diluted 1:1 with buffer D′, applied on a FPLC MonoS column and eluted with a linear gradient of KCl (100 mM–1 M). The pooled active fractions were diluted 1:1 with buffer D′ without KCl and after centrifugation the pool was loaded onto an oligo-affinity column containing the polymerized oligonucleotide alb2 (Kadonaga and Tjian, 1986). This column was eluted with a linear gradient of KCl and active fractions were pooled and diluted again for subsequent cycles on the same oligo-affinity column.

Labelling of oligonucleotides

To generate double-stranded oligonucleotides, equal amounts of complementary single-stranded oligonucleotides at a concentration of 1 $\mu g/\mu l$ were heated for 1 min at 90°C and left to cool at room temperature. The concentration was then adjusted to 1 pmol/μl. For the labelling, 1 pmol of double-stranded oligonucleotide was incubated in the presence of 1 U polynucleotide-kinase, 1 μl of [γ-^{32}P]ATP (6000 Ci/mmol, 10 mCi/ml) in 10 μl of kinase buffer (50 mM Tris pH 7.6, 10 mM $MgCl_2$, 5 mM DTT, 0.1 mM spermidine, 0.1 mM EDTA) at 37°C for 30 min. The reaction mixture was then diluted with 40 μl of H_2O and extracted with 50 μl of phenol/chloroform. After centrifugation the aqueous phase was applied onto a Sephadex G-50 column (250 × 4 mm) equilibrated with TE (10 mM Tris−HCl pH 8, 1 mM EDTA). Fractions of 0.5 ml were collected and counted. Radioactive fractions were pooled and used for the mobility retardation assay.

DNase I footprinting and mobility retardation assays

DNase I footprinting assays were performed according to Lichtsteiner *et al.* (1987) on a −260/+12 fragment from the Alb promoter and on a −340/−120 fragment from the Rbp promoter. For mobility retardation assays, affinity-purified protein was pre-incubated in a 20 μl reaction containing 20 mM Tris pH 7.6, 8% Ficoll, 50 mM KCl, 0.2 mM DTT, 0.3 μg poly(dldC), 3 mM $MgCl_2$ and 3 mM spermidine. After 10 min, 10 000−20 000 c.p.m. of the end-labelled ds oligonucleotide was added and the incubation continued for 10 min at room temperature. Free DNA and DNA−protein complexes were resolved on a 5% polyacrylamide gel in 0.5 × TBE (45 mM Tris-borate, 45 mM boric acid and 2 mM EDTA). After the run the gel was dried onto DE 52 paper and exposed.

Microsequencing and peptide synthesis

About 40 μg of purified protein have been partially sequenced using the method described in Gausepohl *et al.* (1986). Peptide 1 (GIPLESTDGD), peptide 3 (AVKDELLGEK) and peptide 6 (TQDEFHPFIEAL) were synthesized with the addition of a lysine at the amino terminus in a continuous flow instrument constructed and operated as described in Frank and Trosin (1985) and Frank and Gausepohl (1987). Peptide chain assembly was performed by the solid phase method on a 1% crosslinked polystyrene support using fluorenylmethoxycarbonyl amino acids (Carpino and Han, 1972) and *in situ* activation with benzotriazoleyl-oxy-*tris*-(dimethylamino)phosphonium hexafluorophosphate (Castro *et al.*, 1975). The synthetic peptides were purified by reversed-phase HPLC.

Preparation and affinity purification of antibodies

Lyophilized synthetic peptides 1, 3 and 6 were dissolved in PBS at a concentration of 4 $\mu g/\mu l$ and 150 μg were mixed with 450 μg of keyhole limpet haemocyanin (KLH) (Calbiochem). Covalent crosslinking of the peptides to KLH was performed at room temperature by sequential addition at 5-min intervals of five identical aliquots of glutaraldehyde (Polysciences, EM grade, 5% stock solution in water) to a final concentration of 10 mM. The reaction was left for 30 min on ice; the residual amino-reactive groups were blocked by addition of glycine (pH 8.5) to a final concentration of 25 mM. This mixture was used to prepare an emulsion with complete Freund's adjuvant which was injected into the popliteal lymph nodes of the rabbits (150 μg peptide/rabbit). Three weeks later the rabbits were injected again with the same amount of conjugated peptide, mixed with incomplete Freund's adjuvant. After 10 days blood samples were collected and the sera tested for the presence of specific antibody by Western blotting on various preparations of NF1-L. The antibody was immunopurified on a Sepharose 4B-pep-6 column, prepared according to Kreis (1986).

SDS−PAGE and Western blot

SDS−PAGE was performed according to Laemmli (1970) and gels were silver stained according to Wray *et al.* (1981). For the Western blot, the gel was electroblotted onto Immobilon membrane (Millipore) in 20 mM Tris base, 150 mM glycine and 20% methanol at 4°C overnight at 50 V. The membrane was then incubated for 30 min in TBST (10 mM Tris−HCl pH 8, 150 mM NaCl, 0.05 Tween-20 and 5% Carnation nonfat dry milk), 1 h in TBST in the presence of the specific antibody at the desired dilution, washed three times for 10 min each with TBST, 30 min in TBST containing a 1:7500 dilution of anti-rabbit IgG alkaline phosphatase conjugate (Promega), washed three times for 10 min each with TBST. The colour reaction was carried out in 100 mM Tris−HCl pH 9.5, 100 mM NaCl, 5 mM $MgCl_2$ in the presence of the substrates nitro blue tetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl-phosphate (BCIP) at the concentration suggested by the supplier. The reaction was stopped by several washes in water and the membrane was blotted dry. For the denaturation− renaturation experiment shown in Figure 2, ~20 μg of affinity-purified protein was applied to a 12% SDS−PAGE alongside prestained protein markers (BioRad). After the run gel slices were cut and processed as described by Briggs *et al.* (1986).

Preparative band-shift

The experiments were performed essentially as described by Rupp and Sippel (1987). alb2 oligonucleotide was used as the binding site for the binding activity. Radioactive alb2 (5×10^5 c.p.m. in total) was mixed with 60 pmol of cold alb2 and incubated with 15 μl of an active MonoS fraction in a total of 40 μl binding buffer containing 7.5 μg poly(dldC). In the control samples the oligonucleotide or the proteins were omitted. The samples were incubated at room temperature for 20 min and resolved on a 5% polyacrylamide gel. After the run the gel was exposed wet at 4°C for 2 h. A gel slice containing the shifted band together with the adjacent region of the control lanes was excised and applied without any further treatment on the stacking gel of a 12% SDS−gel. After the run the gel was silver stained. For immunostaining, the gel was electroblotted on Immobilon membrane (Millipore) as described above with the following modification. To collect the DNA during the transfer a sheet of DE 52 paper was layered behind the Immobilon membrane. After the transfer the DE 52 paper was dried and exposed while the membrane was immunostained as described above.

Library screening and DNA sequencing

Oligonucleotide 61 (48 mer) was designed on the basis of the peptide 6 as a unique probe using the criteria suggested by Lathe (1985). This ^{32}P-labelled oligonucleotide was used to screen a rat-liver cDNA library in λgt11 (Clontech Laboratories, Inc.) at a concentration of 1 pmol/ml in the following hybridization mixture: 6 × SSC (1 × SSC is 150 mM NaCl and 15 mM sodium citrate), 1 × Denhardt's, 0.05% sodium pyrophosphate and 20 μg/ml tRNA. Hybridization of the filters was carried out at 42°C for at least 12 h, washed twice at 42°C and once at 60°C for 15 min each with 6 × SSC, 0.05% sodium pyrophosphate and exposed at −70°C with intensifying screens. The positive plaques were rescreened with degenerated oligonucleotides corresponding to part of peptide 5 (17 mer containing 64 degeneration). The filters were hybridized at 37°C and washed at 46°C. Three independent clones hybridizing to both oligonucleotides were isolated from 8×10^5 plaques. The cDNA insert of one of these was used to screen the same library again and four more positives were found. Lambda DNA was prepared and sequenced directly according to Manfioletti and Schneider (1988).

Preparation of poly(A)$^+$ RNA

Total RNA from rat tissues was prepared according to Chomczynski and Sacchi (1987). Dissected tissues were homogenized in the presence of D solution (4 M guanidium thiocyanate, 25 mM sodium citrate and 0.1 M 2-mercaptoethanol) and extracted with water-saturated phenol (acid phenol) and chloroform/isoamylalcohol in the presence of 0.2 M sodium acetate pH 4. Two volumes of ethanol were added to the aqueous phase, the RNA precipitates were collected by centrifugation, resuspended in D solution and precipitated again. The pellet was finally resuspended in 0.5% SDS.

To prepare poly(A)-tailed RNA, NaCl to 0.5 M final concentration and 50−100 mg oligo(dT) cellulose were added to the total RNA solution. The suspension was incubated at room temperature on a shaking platform for 8−10 h. Oligo(dT) cellulose was pelleted by centrifugation at 2500 r.p.m. for 30 s, the supernatant was removed and the oligo(dT) cellulose was washed and pelleted again in binding buffer (0.5 M NaCl, 20 mM Tris pH 7.4, 10 mM EDTA, 0.2% SDS). The oligo(dT) cellulose was resuspended in 10 ml of binding buffer, poured into a disposable plastic column and washed wtih ~20 ml of washing buffer (0.1 M Nacl, 10 mM Tris pH 7.4, 0.2% SDS). The poly(A)-tailed RNA was eluted with 1 mM Tris pH 7.8, 1 mM EDTA and 0.2% SDS, precipitated twice with 2 vol ethanol in the presence of 150 mM NaAc pH 5.5, resuspended in TE with 0.2% SDS and stored at −70°C.

Northern blot

About 3 μg of poly(A)-tailed RNA per tissue was applied onto a 1% agarose 2.2 M formaldehyde gel. After the run the gel was soaked for 60 min in 20 × SSC and blotted onto a nylon membrane (GeneScreen) for 12−16 h in the cold room. The filter was dried and UV crosslinked. Prehybridization was carried out in 15 ml of hybridization solution (0.5 M NaH_2PO_4 pH 7.2, 7% SDS and 1 mM EDTA) for 5−15 min at 65°C. Probes labelled as described in Feinberg and Vogelstein (1983) were mixed with 7 ml hybridization solution and left hybridizing with the filter for 20−24 h at 65°C. The filter was washed twice with 1% SDS, 40 mM NaH_2PO_4 pH 7.2 at 65°C, twice with 0.1 × SSC, 0.1% SDS at 65°C and exposed at −70°C with intensifying screen.

Acknowledgements

We are grateful to Lennart Philipson and our colleagues in the Department of Gene Structure and Regulation for critically reading the manuscript. We thank R.Tjian and C.Santoro for communicating their results before publication. We acknowledge the invaluable collaboration of R.Frank and H.Gausepohl, for the determination of the amino acid sequence of the tryptic digest of NF1-L and the synthesis of the peptides, and of Brian Sproat and Philippe Neuner for the synthesis of oligonucleotides. Part of this work was supported by EEC, BAP-grant 115-D. The part of the work done in Italy was supported by P.F., ONCOLOGIA, CNR, Rome, Italy. F.G. was supported by a grant from EEC (one year) and a grant from Boeringher-Ingelheim.

References

Babiss,L.E., Herbst,R.S., Bennet,A.L. and Darnell,J.E., Jr (1987) *Genes Dev.*, **1**, 256–267.
Banerji,J., Olson,L. and Schaffner,W. (1983) *Cell*, **33**, 729–740.
Borgmeyer,U., Nowock,J. and Sippel,A.E. (1984) *Nucleic Acids Res.*, **12**, 4295–4311.
Briggs,M.R., Kadonaga,J.T., Bell,S.P. and Tjian,R. (1986) *Science*, **234**, 47–52.
Buchman,A.R., Kimmerly,W.J., Rine,J. and Kornberg,R.D. (1988) *Mol. Cell. Biol.*, **8**, 210–225.
Carpino,L.A. and Han,G.Y. (1972) *J. Org. Chem.*, **37**, 3404–3409.
Castro,J.R., Dormoy,G., Evin,G. and Selve,C. (1975) *Tetrahedron Lett.*, **14**, 1219–1223.
Cereghini,S., Raymondjean,M., Carranca,A.J., Herbomel,P. and Yaniv,M. (1987) *Cell*, **50**, 627–638.
Chomczynski,P. and Sacchi,N. (1987) *Anal. Biochem.*, **162**, 156–159.
Ciliberto,G., Dente,L. and Cortese,R. (1985) *Cell*, **41**, 531–540.
Ciliberto,G., Arcone,R., Wagner,E.F. and Rüther,U. (1987) *EMBO J.*, **6**, 4017–4022.
Colantuoni,V., Pirozzi,A., Blance,C. and Cortese,R. (1987) *EMBO J.*, **6**, 631–636.
Courtois,G., Morgan,J.G., Campbell,L.A., Fourel,G. and Crabtree,G.R. (1987) *Science*, **238**, 688–692.
Dente,L., Tripodi,M., Rüther,U., Wagner,E. and Cortese,R. (1988) *Genes Dev.*, **2**, 259–266.
De Simone,E., Ciliberto,G., Hardon,E., Paonessa,G., Palla,F., Lundberg,L. and Cortese,R. (1987) *EMBO J.*, **6**, 2759–2766.
Dignam,J.D., Lebovitz,R.M. and Roeder,R.G. (1983) *Nucleic Acids Res.*, **11**, 1475–1489.
D'Onofrio,C., Colantuoni,V. and Cortese,R. (1985) *EMBO J.*, **4**, 1981–1989.
Feinberg,A.P. and Vogelstein,B. (1983) *Anal. Biochem.*, **132**, 6–13.
Frank,R. and Gausepohl,H. (1987) In Tschesche,H. (ed.), *Modern Methods in Protein Chemistry*. Walter de Gruyter, Berlin, Vol. 3, pp. 41–60.
Frank,R. and Trosin,M.. (1985) In Tschesche,H. (ed.), *Modern Methods in Protein Chemistry*. Walter de Gruyter, Berlin, Vol. 2, pp. 287–302.
Gausepohl,H., Trosin,M. and Frank,R. (1986) In Wittmann-Liebold,B., Salnikow,J. and Erdmann,V.A. (eds), *Advanced Methods in Protein Microsequence Analysis*. Springer Verlag, Berlin, pp. 149–160.
Gillies,S.D., Morrison,S.L., Oi,V.T. and Tonegawa,S. (1983) *Cell*, **33**, 717–728.
Godbout,R., Ingram,R. and Tilghman,S.M. (1986) *Mol. Cell. Biol.*, **6**, 477–487.
Godbout,R., Ingram,R. and Tilghman,S.M. (1988) *Mol. Cell. Biol.*, **8**, 1169–1178.
Gorski,K., Carneiro,M. and Schibler,U. (1986) *Cell*, **47**, 767–776.
Grayson,D.R., Costa,R.H., Xanthopoulos,K.G. and Darnell,J.E., Jr (1988a) *Mol. Cell. Biol.*, **8**, 1055–1066.
Grayson,D.R., Costa,R.H., Xanthopoulos,K.G. and Darnell,J.E., Jr (1988b) *Science*, **239**, 786–788.
Hardon,E.M., Frain,M., Paonessa,G. and Cortese,R. (1988) *EMBO J.*, **7**, 1711–1719.
Heard,J., Herbomel,P., Ott,M., Mottura-Rollier,A., Weiss,M. and Yaniv,M. (1987) *Mol. Cell. Biol.*, **7**, 2425–2434.
Jones,C.N., Rigby,P.W. and Ziff,E.B. (1988) *Genes Dev.*, **2**, 267–281.
Jones,K.A., Kadonaga,J.T., Rosenfeld,P.J., Kelly,T.J. and Tjian,R. (1987) *Cell*, **48**, 78–89.
Kadonaga,J.T. and Tjian,R. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 5889–5893.
Kelsey,G.D., Povey,S., Bygrave,A.E. and Lovell-Badge,R.H. (1987) *Genes Dev.*, **1**, 161–171.
Kreis,T.E. (1986) *EMBO J.*, **5**, 931–941.
Kugler,W., Wagner,U. and Ryffel,G.U. (1988) *Nucleic Acids Res.*, **16**, 3165–3174.
Laemmli,U.K. (1970) *Nature*, **277**, 680–685.
Lambert,P.F., Spalholz,B.A. and Howley,P.M. (1987) *Cell*, **50**, 69–78.
Lathe,R. (1985) *J. Mol. Biol.*, **183**, 1–12.
Lee,W., Mitchell,P. and Tjian,R. (1987) *Cell*, **49**, 741–752.
Leegwater,P.A., van Driel,W. and van der Vliet,P.C. (1985) *EMBO J.*, **4**, 1515–1521.
Lichtsteiner,S., Wuarin,J. and Schibler,U. (1987) *Cell*, **51**, 963–973.
Lillie,J.W., Green,M. and Green,M.R. (1986) *Cell*, **46**, 1043–1051.
Manfioletti,G. and Schneider,C. (1988) *Nucleic Acids Res.*, **16**, 2873–2884.
Maniatis,T., Goodburn,S. and Fisher,J.A. (1987) *Science*, **236**, 1237–1244.
Monaci,P., Nicosia,A. and Cortese,R. (1988) *EMBO J.*, **7**, 2075–2087.
Muglia,L. and Rothman-Denes,L.B. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 7653–7657.
Nagata,K., Guggenheimer,R.A. and Hurwitz,J. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 6177–6181.
Neuberger,M.S. (1983) *EMBO J.*, **2**, 1373–1378.
Oliviero,S., Morrone,G. and Cortese,R. (1987) *EMBO J.*, **6**, 1905–1912.
Ott,M.O., Sperling,L., Herbomel,P., Yaniv,M. and Weiss,M.C. (1984) *EMBO J.*, **3**, 2505–2510.
Pinkert,C.A., Ornitz,D.M., Brinster,R.L. and Palmiter,R.D. (1987) *Genes Dev.*, **1**, 268–276.
Queen,C. and Baltimore,D. (1983) *Cell*, **33**, 741–748.
Raymondjean,M., Cereghini,S and Yaniv,M. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 757–761.
Rosenfeld,P.J. and Kelly,T.J. (1986) *J. Biol. Chem.*, **261**, 1398–1408.
Rupp,R.A.W. and Sippel,A.E. (1987) *Nucleic Acids Res.*, **15**, 9707–9726.
Rüther,U., Tripodi,M., Cortese,R. and Wagner,E.F. (1987) *Nucleic Acids Res.*, **15**, 7519–7529.
Santoro,C., Mermod,N., Andrews,P.C. and Tjian,R. (1988) *Nature*, **334**, 218–224.
Schwartz,R.M. and Dayhoff,M.O. (1978) In Dayhoff,M.O. (ed.), *Atlas of Protein Sequence and Structure*. National Biomedical Research Foundation, Washington, DC, Vol. 5, pp. 353–358.
Shaul,Y., Ruttner,W.J. and Laub,O. (1985) *EMBO J.*, **4**, 427–430.
Shaul,Y., Ben-Levy,R. and De-Medina,T. (1986) *EMBO J.*, **5**, 1967–1971.
Shore,D. and Nasmyth,K. (1987) *Cell*, **51**, 721–732.
Walker,M.D., Edlund,T., Boulet,A.M. and Rutter,W.J. (1983) *Nature*, **306**, 557.
Wray,W., Boulikas,T., Wray,V.P. and Hancock,R. (1981) *Anal. Biochem.*, **118**, 197–203.

Received on June 13, 1988; revised on July 6, 1988