Expression, DNA-binding specificity and transcriptional regulation of nuclear factor 1 family proteins from rat

Shigehiro OSADA, Tsukasa MATSUBARA, Shoko DAIMON, Yoshie TERAZU, Mingxu XU, Tsutomu NISHIHARA and Masayoshi IMAGAWA[1]

Laboratory of Environmental Biochemistry, Graduate School of Pharmaceutical Sciences, Osaka University, 1–6 Yamada-Oka, Suita, Osaka 565-0871, Japan

Nuclear factor 1 (NF1) family proteins, which are encoded by four different genes (NF1-A, NF1-B, NF1-C and NF1-X), bind to the palindromic sequence and regulate the expression of many viral and cellular genes. We have previously purified NF1-A and NF1-B from rat liver as factors that bind to the silencer in the glutathione transferase P gene, and have also reported the repression domain of NF1-A. In the present study we cloned five cDNA species (NF1-B1, NF1-B2, NF1-B3, NF1-C2 and NF1-X1) and compared their expression profiles and the affinity and specificity of the DNA binding of these NF1 family members. By Northern blot analysis, we found that the expression profiles of the NF1s are indistinguishable in the various tissues of the rat. The DNA-binding affinities of NF1-A and NF1-X are higher than those of NF1-B and NF1-C, whereas all four NF1 proteins showed the same DNA-binding specificity. Transfection analyses revealed that the function of NF1-B on the transcriptional regulation differed between NF1-B isoforms and was affected by the factor(s) that bind to the promoter regions. In addition, we identified the transcriptional regulatory domain of NF1-B, which is enriched with proline and serine residues.

Key words: gene expression, transcriptional activation domain, transcription factor.

INTRODUCTION

Many transcription factors have been found to be members of highly related multifactor families; their specificity of action must therefore be addressed to ascertain their respective functions [1]. Transcriptional factors consist of two functionally distinct domains: a DNA-binding domain and a transcriptional regulatory domain [2].

Nuclear factor 1 (NF1) is one of the transcription factor family proteins. The molecular cloning of NF1 cDNA species from several species made it clear that NF1 proteins were produced not only by alternative splicing but also by four different genes (NF1-A, NF1-B, NF1-C and NF1-X) [3–9]. The conserved approx. 230-amino-acid residue N-terminal region of all NF1 proteins is important for the DNA binding, dimerization and stimulation of adenovirus DNA replication. The C-terminal regions of NF1 proteins, in contrast, are diverse and required for the regulation of transcription [9–12].

Human NF1-C was initially purified from HeLa nuclear extracts as a factor that promotes adenovirus DNA replication. Expression and functional analyses of the cloned NF1-C cDNA revealed that NF1-C binds to the GCCAAT sequence and activates adenovirus DNA replication and transcription [3]. Subsequently, rat NF1-A was purified and cloned from the liver as a factor that recognizes the NF1 site on the albumin gene [4]. NF1-B and NF1-X (from hamster liver), which bind to the promoter for the 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) reductase gene, have been purified [4]. We recently purified the factors that bind to the silencer element in the rat glutathione transferase P (GST-P) gene. The amino acid sequences of the purified factors were identical with those of NF1-A and NF1-B [9]. Although the four NF1 proteins listed above were purified from several species, their expression profiles are unclear. Moreover, whereas all the NF1 proteins recognize the origin of the adenovirus replication, their affinities and specificities of DNA binding to other NF1 sites are still unknown.

The C-terminal regions of NF1 proteins are diverse and required for the regulation of transcription. We previously identified a rat NF1-A repression region that is divided into two 100-residue domains, one enriched with serine and glycine residues and the other with proline and serine residues [13]. The proline-rich activation domain was identified in the C-terminal region of human NF1-C and selectively interacts with TFIIB, TATA-binding protein (TBP), TBP-associated factor 55 and histone H3 [10,14–17]. Roulet et al. [11] reported that the C-terminal region of *Xenopus* NF1-X activates transcription by using an activation domain that is subdivided into two domains, each containing independent activation properties in chimaeric proteins. NF1-B has been purified as a silencer or promoter-binding factor, but the C-terminal region of NF1-B has not yet been characterized.

In the present study, we cloned all four NF1 cDNA species of the rat and found that their expressions are indistinguishable in rat tissues. The DNA-binding affinities of NF1-A and NF1-X were revealed to be higher than those of the other NF1 proteins, although all four NF1 proteins showed the same DNA-binding

Abbreviations used: C/EBP, CCAAT-enhancer-binding protein; GST-P, glutathione transferase P; HMG-CoA, 3-hydroxy-3-methylglutaryl-CoA; NF1, nuclear factor 1; RT–PCR, reverse-transcriptase-mediated PCR; SV40, simian virus 40; TBP, TATA-binding protein.

[1] To whom correspondence should be addressed (e-mail imagawa@phs.osaka-u.ac.jp.).

The nucleotide sequence data reported will appear in DDBJ, EMBL and GenBank Nucleotide Sequence Databases under the following accession numbers: NF1-B1, AB012230; NF1-B2, AB012231; NF1-B3, AB012232; NF1-C2, AB012233; C-terminal portion of NF1-X1, AB012234; and N-terminal portion of NF1-X1, AB012235.

© 1999 Biochemical Society

specificity. We also observed that the function for the transcriptional regulation of NF1-B is affected by the factor(s) that bind to the promoter regions. Using a minimal promoter, we identified the NF1-B transcriptional activation domain, which is enriched with proline and serine residues.

MATERIALS AND METHODS

Library screening and DNA sequencing

The specific probes for NF1 family cDNA species were amplified by a reverse-transcriptase-mediated PCR (RT–PCR) analysis by using rat liver RNA as a template with three sets of two primers corresponding to hamster NF1-B (position 757–1570 bp), human NF1-C (position 723–1464 bp) and hamster NF1-X (position 713–1326 bp) [3,5]. The amplified PCR products were cloned into a pBluescript vector and sequenced with automated DNA sequencers [DSQ-1000 (Shimadzu Ltd., Kyoto, Japan) and ABI PRISM 310 (Perkin-Elmer, Foster City, CA, U.S.A.)]. The sequences revealed that the PCR products are the counterparts of the rat NF1 cDNA species. ^{32}P-labelled rat NF1-B, NF1-C, and NF1-X cDNA fragments were used as probes for screening 6×10^5 phages of a rat liver λgt10 cDNA library kindly provided by Dr. N. Miura (Akita University, Akita, Japan). We obtained three NF1-B clones, three NF1-C clones and two NF1-X clones; each insert of all isolated clones was subcloned into a pBluescript vector. The nucleotide sequence was determined with automated DNA sequencers (DSQ 1000 and ABI PRISM 310) and the dideoxy method by using ^{32}P [18].

RT–PCR analysis

The reverse transcriptase reaction was performed for 30 min at 55 °C with 1.5 μg of total RNA as a template. NF1-B isoforms were detected by the primers corresponding to position 1160–1688 bp of NF1-B1. All three NF1-B isoforms could be detected by this primer set. The sizes of the products of NF1-B1, NF1-B2 and NF1-B3 were 538, 449 and 227 bp respectively. PCR amplification (94 °C for 1 min, 61 °C for 1 min and 72 °C for 1 min) was performed for 30 cycles. The reaction products were separated on a 6% (w/v) polyacrylamide gel and stained with ethidium bromide. The positive fragments were recovered, blunted, cloned into the *Sma*I site of the pBluescript vector and sequenced by an automated DNA sequencer (ABI PRISM 310).

Northern blot analysis

Poly(A)$^+$ RNA was isolated from total RNA by using Oligotex-dT30 (Daiichi Pure Chemicals, Tokyo, Japan). mRNA (2 μg) was subjected to electrophoresis on a 1% (w/v) agarose gel containing 2% (v/v) formaldehyde, then transferred to a nitrocellulose filter. For the detection of NF1-A, the region corresponding to 949–1530 bp in NF1-A cDNA was used as a probe. The other NF1 isoforms were detected by the specific probes used for the screening.

Plasmid constructs

For the expression of the DNA-binding domains of NF1 family proteins as GST fusion proteins, the fragments corresponding to amino acid residues 10–244 (NF1-A), 11–239 (NF1-B), 10–246 (NF1-C), and 10–245 (NF1-X) were amplified by PCR and then cloned into the pGEX-2T (Pharmacia, Uppsala, Sweden).

GAL4-A1(145–509), also previously termed GAL4-A1 [9], contains the C-terminal region of rat NF1-A. GAL4-B2(239–494) and GAL4-B3(239–420) were made by digesting NF1-B cDNA species with *Hin*cII and *Xba*I, then subcloning into the *Bam*HI fill-in and the *Xba*I site of pSG424 [19]. GAL4-B1(239–561) was obtained by inserting an *Eco*RV–*Kpn*I fragment from NF1-B1 cDNA into the *Eco*RV–*Kpn*I site of GAL4-B2. GAL4-C1(390–499) and GAL4-C2(399–439) were constructed by inserting RT–PCR-amplified fragments corresponding to amino acid residues 390–499 of human NF1-C1 and 399–439 of rat NF1-C2 respectively. The series of NF1-B mutants were constructed in the pSG424 vector encoding the DNA-binding domain of GAL4 by inserting the appropriate position of NF1-B cDNA species and are termed GAL4-B(x–y), where x and y indicate the start and end residues respectively of the NF1-B proteins inserted.

$5\times$GAL4–91GST–luciferase, $5\times$GAL4–167C/EBPδ–luciferase (in which C/EBP stands for CCAAT-enhancer-binding protein) and $5\times$GAL4–MTIIA–luciferase have been described previously [9,13]. For the construction of $5\times$GAL4–SV40–luciferase, the fragment including the pentamer of the GAL4 binding site was cloned into the *Kpn*I–*Bgl*II site of the luciferase vector PGV-P (Toyo Ink, Tokyo, Japan) containing the simian virus 40 (SV40) promoter. $5\times$GAL4–luciferase vector was generated by removing the promoter region from $5\times$GAL4–MTIIA–luciferase by digestion with *Hin*dIII. To make the constructs of $5\times$GAL4–TATA (GST-P)–luciferase and $5\times$GAL4–TATA (MTIIA)–luciferase, the fragment from bp -33 to $+57$ and -43 to $+69$ in the minimum promoter of the GST-P and MTIIA genes respectively were cloned into the $5\times$GAL4–luciferase vector. $5\times$GAL4–TATA (E1b)–luciferase was constructed by inserting the fragment including the pentamer of the GAL4-binding site and the E1b minimum promoter derived from pG5CAT (Clontech, Palo Alto, CA, U.S.A.) into the *Xho*I–*Hin*dIII site of the luciferase vector PGV-B (Toyo Ink). The sequences of all fragments generated by PCR were checked with an automated DNA sequencer (ABI PRISM 310).

Expression and purification of the DNA-binding domains of NF1 family proteins

The NF1 DNA-binding domain expression plasmids were transformed into JM109 or BL21. The transformant was grown to a D_{600} of 0.6 at 30 °C in Luria–Bertani medium; isopropyl β-D-thiogalactoside was then added to a final concentration of 0.1 mM. The cells were then allowed to grow for 16 h at 16 °C, harvested by centrifugation and suspended in 10% of the original culture volume in 0.5 M HM buffer [25 mM Hepes (pH 7.9)/12.5 mM $MgCl_2$/1 mM dithiothreitol/20% (v/v) glycerol/1 μM pepstatin/4 μM leupeptin/1 mM 4-(2-aminoethyl) benzenesulphonyl fluoride hydrochloride, with KCl as indicated], then disrupted by sonication. After centrifugation at 418000 ***g*** (98000 rev./min) for 0.5 h, the supernatant added to 0.1% Nonidet P40 was mixed with glutathione–Sepharose beads and rotated at 4 °C for 30 min. The GST fusion proteins were eluted with 0.5 M HM buffer containing 20 mM glutathione.

Gel-shift assay

The sequences of oligonucleotides for the gel-shift assay were as follows (only upper strands are shown): adenovirus replication origin, 5′-CTAGCTATTTTGGATTGAAGCCAATATG-3′; GST-P, 5′-CTAGTTTCTTGGAGCAGGACCCAAAAAT-3′; α-globin, 5′-CTAGCGGGCTCCGCGCCAGCCAATGAGC-3′; HMG-CoA reductase, 5′-CTAGTGATGCTGGAACTCG-ACCAGCTAT-3′.

The protein fraction (8.25 μl) containing 2 μg of BSA and a constant amount of NF1 proteins (2 ng of NF1-A, 6 ng of NF1-

© 1999 Biochemical Society

B, 25 ng of NF1-C or 8 ng of NF1-X) was mixed with 4.25 μl of 20 mM Tris/HCl, pH 7.5, 10% (v/v) glycerol, 2 mM dithiothreitol, 10 mM EDTA and labelled probe. Probe DNA was added to a final concentration of 0.25–8.0 nM. Each reaction mixture was incubated for 30 min at room temperature, loaded on a 6% (w/v) non-denaturing polyacrylamide gel, subjected to electrophoresis at 150 V for 1 h and fixed with 10% (v/v) methanol and 10% (v/v) acetic acid. The radioactivity corresponding to each band was measured by a bioimage analyser BAS 2000 (Fuji Film, Tokyo, JApan).

Cell culture, transfection and luciferase assay

HeLa cells were cultured in minimal essential medium supplemented with 10% (v/v) fetal bovine serum, then passaged by treatment with trypsin at confluence. Transfection was performed by the calcium phosphate co-precipitation technique [20], with 2.0 μg of the luciferase reporter plasmid, 0.5 μg of the effector plasmid, 2.0 μg of pBluescript and 0.5 μg of pRSVGAL, a eukaryotic expression vector containing the *Escherichia coli* β-galactosidase (*lacZ*) structural gene controlled by the Rous sarcoma virus long terminal repeat, as an internal control. Luciferase activities were measured with the use of Pikka Gene (Toyo Ink). All transfection experiments were performed three to eight times with two or three different preparations of DNA, and the relative luciferase activity was derived from the mean values of the results. The activity of β-galactosidase was assayed as described [21].

Preparation of nuclear extracts and Western blot analysis

The preparation of nuclear extracts from the transfected HeLa cells and the Western blot analyses were performed as described previously [22]. The nuclear extracts were subjected to SDS/PAGE [10–20% (w/v) gradient gel], transferred to nitrocellulose and immunoblotted by the anti-(GAL4 DNA-binding domain) antibody. Anti-(GAL4 DNA-binding domain) antibody, corresponding to residues 1–147, was purchased from Santa Cruz Biotechnology (Santa Cruz, CA, U.S.A.).

RESULTS

Isolation of cDNA clones encoding rat NF1 family proteins

We previously cloned rat NF1-A cDNA from a λgt10 rat liver library. To obtain the other isoforms of NF1, the specific probes corresponding to the C-terminal regions of NF1 were obtained by RT–PCR (results not shown). The sequences of the three amplified fragments were similar to hamster NF1-B, human NF1-C2 and hamster NF1-X respectively. By using these fragments as probes, we obtained three NF1-B clones, three NF1-C clones and two NF1-X clones from the λgt10 rat liver library, and determined the sequences of each insert of all isolated clones (Figure 1). Unfortunately, none of these eight clones had a complete open reading frame. However, it is likely that three types of NF1-B cDNA (NF1-B1, NF1-B2, NF1-B3) were isolated, and both NF1-B2 and NF1-B3 are deleted forms of the C-terminal portion of NF1-B1. These three clones contained the fragment corresponding to residue 193 to the stop codon of NF1-B2, residue 137 to the stop codon of NF1-B3, and the start codon to residue 437 of NF1-B1. The complete NF1-B1 sequence is shown in Figure 1(A) as the combined sequence with that derived from RT–PCR for the C-terminal region. Because the 5′ sequences of the isolated regions of NF1-B2 and NF1-B3 are exactly the same as that of NF1-B1, we assumed the N-terminal portion of these three NF1-Bs to be identical in later experiments. The combined sequences of the overlapping NF1-C clones and NF1-X clones contained 432 residues of NF1-C2 and 441 residues of NF1-X1 respectively.

The N-terminal regions (181 residues) of NF1 proteins were highly conserved. The four cysteine residues that are essential for DNA binding and the charged amino acids in the region predicted by two algorithms to have a high α-helix probability, termed the lysine helix, are also conserved (Figure 1A) [23,24].

From a sequence comparison of the three NF1-B clones, NF1-B2 and NF1-B3 seemed to be isoforms created by alternative splicing. The frame-shifted NF1-B2 and NF1-B3 have a Ser-Trp-Tyr-Leu-Gly motif in the C-terminal region (Figure 1B). To determine the mRNA expressions of these NF1-B cDNA species in rat liver, spleen, kidney, heart, lung, intestine, testis, brain and muscle, we performed RT–PCR with the common primers to three NF1-Bs (see Figure 1B). These primers produce three different lengths of isoforms. A single band of the same size corresponding to NF1-B3 was detected in all tissues (Figure 1C). Sequence analysis revealed that this single band was derived from the NF1-B3 mRNA. The bands derived from NF1-B1 and NF1-B2 are rarely detected in all tissues tested. PCR analysis with this set of primers and NF1-B1, NF1-B2 and NF1-B3 cDNA species inserted in pBluescript as templates showed that the amounts of product of NF1-B3 was approximately twice that of others (results not shown). These results indicate that NF1-B3 might be the dominant form of NF1-Bs in the rat tissues, although the amplification efficiency of PCR for NF1-B3 is slightly greater than that of the other two NF1-Bs, probably owing to the shorter length of the fragment.

The NF1-C and NF1-X clones obtained were the rat counterparts of human NF1-C2 and hamster NF1-X respectively. We have not isolated the other splicing isoforms of NF1-C and NF1-X [3,5].

Expression profiles of NF1 mRNA species

The tissue distributions of the NF1 mRNA species were examined by Northern blot analysis (Figure 2). For each of NF1-A and NF1-C, two bands, at 11.5 and 6 kb and at 8.9 and 4.7 kb respectively, were detected. The NF1-B and NF1-X mRNA species each expressed one major transcript, at 8.7 and 6.6 kb respectively. All of the NF1 mRNA species were detected ubiquitously in various tissues, although the level of NF1 was higher in the lungs and lower in the liver. The transcript of NF1-B in the liver was detected when the film was exposed for 4 days (results not shown).

DNA-binding specificities and affinities of NF1 family members

The DNA-binding-site preferences for the NF1 family proteins are unclear, although all four of the NF1 proteins bind to the TGGCA sequence and the origin of the adenovirus replication [25]. Different NF1 proteins were purified as factors that recognize the *cis*-elements containing TGG [26]. There is a possibility that each of the NF1 proteins recognizes a subtly different sequence and has a different affinity. We previously identified the recognition sequence for NF1-A by the PCR-mediated random site selection method, which is useful for defining the DNA-binding site of transcription factors in an unbiased manner [27]. However, the recognition sequences of the NF1 family proteins are so similar that it is difficult to distinguish between them by this method. We performed a gel-shift assay to determine the dissociation constants for each protein for the four NF1 sites. We previously cloned the rat NF1-A gene and showed that the second exon starts at the tenth amino acid [28]. Residues 10–244 of NF1-A (and the similar regions of the other NF1

© 1999 Biochemical Society

A

```
NF1-A1   1:-MYSPLCLTQDEFHPFIEALLPHVRAFAYTWFNLQARKRKYFKKHEKRMSKEEERAVKDE  59
NF1-B1   1:MMYSPICLTQDEFHPFIEALLPHVRAIAYTWFNLQARKRKYFKKHEKRMSKDEERAVKDE  60
NF1-C2   8:-------LTQDEFHPFIEALLPHVRAFAYTWFNLQARKRKYFKKHEKRMSKDEERAVKDE  60
NF1-X1   1:-MYSPYCLTQDEFHPFIEALLPHVRAFSYTWFNLQARKRKYFKKHEKRMSKDEERAVKDE  59
           **** ******************* ********************* ********

NF1-A1  60:LLSEKPEVKQKWASRLLAKLRKDIRPEYREDFVLTVTGKKPPCCVLSNPDQKGKMRRIDC 119
NF1-B1  61:LLSEKPEIKQKWASRLLAKLRKDIRQEYREDFVLTVTGKKHPCCVLSNPDQKGKIRRIDC 120
NF1-C2  61:LLGEKAEVKQKWASRLLAKLRKDIRPECREDFVLAVTGKKAPGCVLSNPDQKGKMRRIDC 120
NF1-X1  60:LLGEKPEIKQKWASRLLAKLRKDIRPEFREDFVLTITGKKPPCCVLSNPDQKGKIRRIDC 119
           ** ** * ***************** * ******  **** * *********** *****

NF1-A1 120:LRQADKVWRLDLVMVILFKGIPLESTDGERLVKSPQCSNPGLCVQPHHIGVSVKELDLYL 179
NF1-B1 120:LRQADKVWRLDLVMVILFKGIPLESTDGERLMKSPHCTNPALCVQPHHITVSVKELDLFL 180
NF1-C2 121:LRQADKVWRLDLVMVILFKGIPLESTDGERLVKAAACAHPVLCVQPHHIGVAVKELDLYL 180
NF1-X1 120:LRQADKVWRLDLVMVILFKGIPLESTDGERLYKSPQCSNPGLCVQPHHIGVTIKELDLYL 179
           ****************************** *   *  * ******** *  ***** *

NF1-A1 180:AYFVHAADSSQSESPSQPSDADI-KDQPEN-GHLGFQDSFVTSGVFSVTELVRVSQTPIA 237
NF1-B1 181:AYYVQEQDSGQSGSPSHN-DPAK-NP-PGYLE--D---SFVKSGVFNVSELVRVSRTPIT 232
NF1-C2 181:AYFVRERDAEQSSSPRTGVGSDQEDSKPITLDTTDFQESFVTSGVFSVTELIQVSRTPVV 240
NF1-X1 180:AYFVHTPESGQSDSSNQQGDADI-KPLP-N-GHLSFQDCFVTSGVWNVTELVRVSQTPVA 236
           ** *      ** *             *             ** ***  * **  ** **

NF1-A1 238:AGTGPNFSLSDLESSSYYSMSPGAM-RRSLPSTSSTSSTKRLKSVED-EMDSPGEEPFYT 295
NF1-B1 233:QGTGVNFPIGEIPSQPYYHDMNSGVNLQRSL-S-SPPSSKRPKTISIDENMEPSPTG--D 288
NF1-C2 241:TGTGPNFSLGELQGHLAYDLNPASAGMRRTLPSTSSSGSKRHKSGSMEEDVDTSPGG--D 298
NF1-X1 237:TASGPNFSLADLESPSYYNINQVTLGRRSITSPPSTSTTKRPKSIDDSEMESPVDDVFYP 296
             * **        *            *     ** *       *

NF1-A1 296:GQGRSPGSG-SQSSGW-HEVEPGMPSPTTLKKSEKSGFSSPSPSQTSSLGTAFTQHHRP- 352
NF1-B1 289:FY-PSPNSPAAGSRTW-HERDQDMSSPTTMKKPEKPLFSSTSPQDSSPRLSTFPQHHHPG 346
NF1-C2 299:YY-TSPNSPTSSSRNWTEDMEGGISSPVKKTEMDKSPFNSPSPQD-SPRLSSFTQHHRPV 356
NF1-X1 297:GTGRSPAAGSSQSSGW-PN-DVD-AGPASLKKSGKLDFCSALSSQGSSPRMAFTHHPLP- 352
              **      *  *         *        *  * *        *     *   *   *

NF1-A1 353:VITGPR--AS-PHA-TPSTLHFPTSPIIQQPGP-YFSHPAIRY--HP--QETLKEFV--- 400
NF1-B1 347:IPGVAHSVISTRTPPPPSPLPFPTQAILPPAPSSYFSHPTIRYPPHLNPQDTLKNYVPSY 406
NF1-C2 357:I-AV-HSGI-ARSPHPSSALHFPATPILPQTASTYFPHTAIRYPPHLNPQDPLKDLV-SL 412
NF1-X1 353:VLAGVRP-GS-PRA-TASALHFPSTSIIQQSSP-YFTHPTIRYHHHHG-QDSLKEFV--- 404
                          * * **   *        ** *  ***  *   *  **  *

NF1-A1 401:QLVCPDAGQQAGQVGFLNPNGSSQGKVHNP-FLPTPMLPPPPPPPMARPVPLPMPD-TKP 458
NF1-B1 407:DPSSPQTSQPNSSGQVVGKVPGHFTPVLAPSPHPSAVRPVTLTMTDTKPITTSTEAYTAS 466
NF1-C2 413:-ACDPATQQPGPPALRPAR-PLQTVPLWD                                439
NF1-X1 405:QFVCSD-G--SGQAT--GQH--SQ-R-QAP-PLPTG-LSASDPGTATF             441

NF1-A1 459:PTTSTEG-GAASPTSPTY-STPSTSPANRFVSVGPRDPSFVNIPQQT-Q-SW-YLG     509
NF1-B1 467:GTSQANRYVGLSPRDPSFLHQQQLRICDWTMNQNGRH-LYPSTSEDTLGITWQSPGTWAS 525

NF1-B1 526:LVPFQVSNRTPILPANVQNYGLNIIGEPFLQAETSN                         561
```

B

```
NF1-B  caccgttgccatttccaacgcaagctatccttcctccagcaccatcaagttacttttctc 1150
NF1-B1  P  L  P  F  P  T  Q  A  I  L  P  P  A  P  S  S  Y  F  S  H   384
NF1-B2  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
NF1-B3  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -

NF1-B  atccaacaatcagatatcctcctcacctgaatcctcaggatactctgaagaactatgtac 1210
NF1-B1  P  T  I  R  Y  P  P  H  L  N  P  Q  D  T  L  K  N  Y  V  P   404
NF1-B2  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
NF1-B3  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -

NF1-B  cttcttatgacccgtccagtcctcaaacgagccagcctaacagcagtggtcaagtagtag 1270
NF1-B1  S  Y  D  P  S  S  P  Q  T  S  Q  P  N  S  S  G  Q  V  V  G   424
NF1-B2  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
NF1-B3  -  -  -  -  -  -  -  -  -  -  -  /

NF1-B  ggaaagtgcctggccatttcacgcctgtcttggcaccctctccccatcccagtgcagtgc 1330
NF1-B1  K  V  P  G  H  F  T  P  V  L  A  P  S  P  H  P  S  A  V  R   444
NF1-B2  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -

NF1-B  gacctgtgaccctgaccatgacagatactaaacccatcactacatccactgaagcctaca 1390
NF1-B1  P  V  T  L  T  M  T  D  T  K  P  I  T  T  S  T  E  A  Y  T   464
NF1-B2  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -

NF1-B  cagcctcaggcacatctcaagccaatcgatatgtgggactaagcccaagagacccatcct 1450
NF1-B1  A  S  G  T  S  Q  A  N  R  Y  V  G  L  S  P  R  D  P  S  F   484
NF1-B2  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -

NF1-B  tcctgcatcagcaacagctgaggatttgtgactggaccatgaatcaaaacggcaggcatt 1510
NF1-B1  L  H  Q  Q  Q  L  R  I  C  D  W  T  M  N  Q  N  G  R  H  L   504
NF1-B2  -  -  -  -  -  /

NF1-B  tataccccagtaccagtgaggatacattgggaattacttggcaaagtcctggtacctggg 1570
NF1-B1  Y  P  S  T  S  E  D  T  L  G  I  T  W  Q  S  P  G  T  W  A   524
NF1-B2                                                /S  W  Y  L  G   494
NF1-B3                                                /S  W  Y  L  G   420

NF1-B  ctagcttggttccttttcaagtgtcaaataggacacccatcttaccggcaaatgtccaaa 1630
NF1-B1  S  L  V  P  F  Q  V  S  N  R  T  P  I  L  P  A  N  V  Q  N   544
NF1-B2 *
NF1-B3 *

NF1-B  attatggtttgaacataattggagaacctttccttcaagcggagacaagcaactgagggg 1690
NF1-B1  Y  G  L  N  I  I  G  E  P  F  L  Q  A  E  T  S  N  *         561

NF1-B  aaagaaacacaaccatagtgtaagaaaattaggggaaaaacacccaacaaaagaaaagag 1750
```

C

marker liver spleen kidney heart lung intestine testis brain muscle

710 489 404 364 242 190

NF1-B1 NF1-B2 NF1-B3

Figure 1 Deduced amino acid sequences of the rat NF1 proteins and the relative expression of the NF1-B isoforms

(**A**) Sequence alignment of the rat NF1 proteins. Gaps introduced to maximize similarity are indicated by dashes. Asterisks indicate positions that are identical in all four NF1 proteins. The four cysteine residues that are essential for the DNA binding are indicated by a box; the putative lysine helix region is indicated by a shaded box. (**B**) Nucleotide and amino acid sequences of the C-terminal region of rat NF1-B isoforms. Amino acid sequences identical to those above are indicated by dashes. The solidus and the asterisk indicate the deleted portion and the stop codon respectively. The Ser-Trp-Tyr-Leu-Gly motifs in NF1-B2 and NF1-B3 are indicated by boxes. The primers used in RT–PCR for the detection of NF1-B isoforms are underlined. (**C**) Expression levels of mRNA of NF1-B isoforms. To detect the expression levels of NF1-B isoforms in several rat tissues, RT–PCR was performed with common primers for all three NF1-B isoforms. The sizes of the products of NF1-B1, NF1-B2 and NF1-B3 were 538, 449 and 227 bp respectively. pBluescript digested with *Hap*II was used as a size marker. Tissues are indicated above each lane.

proteins) were expressed and used for the gel-shift assay, because multiple splicing isoforms containing the different first exons of NF1 have been cloned [6,29]. A constant amount of NF1 protein was titrated with increasing amounts of a radiolabelled NF1 site derived from the adenovirus replication origin, α-globin, HMG-CoA reductase and GST-P genes. NF1-bound probes were separated from free probes in the binding mixtures by PAGE. The radioactivity of each band was measured by an imaging analyser. These assays revealed that the amount of bound probe approached saturation within the range of input in all assays, with all four NF1 proteins (Figure 3, and results not shown). The results were subjected to a Scatchard analysis, allowing the calculation of dissociation constants. In the lower input range, the amount of bound probe derived from adenovirus replication origin increased more rapidly than that of the NF1 site in the GST-P gene. The level of binding saturation of the adenovirus NF1 site is higher than that of the GST-P NF1 site. The dissociation constant of the GST-P NF1 site shown by the Scatchard analysis slope was approx. 5-fold the corresponding value of the adenovirus.

The dissociation constants for each protein to the four NF1 sites are given in Table 1. The order of affinity among the four NF1 sites in all NF1 proteins was dependent on the similarity to the NF1-A consensus sequence derived from the PCR-mediated

© 1999 Biochemical Society

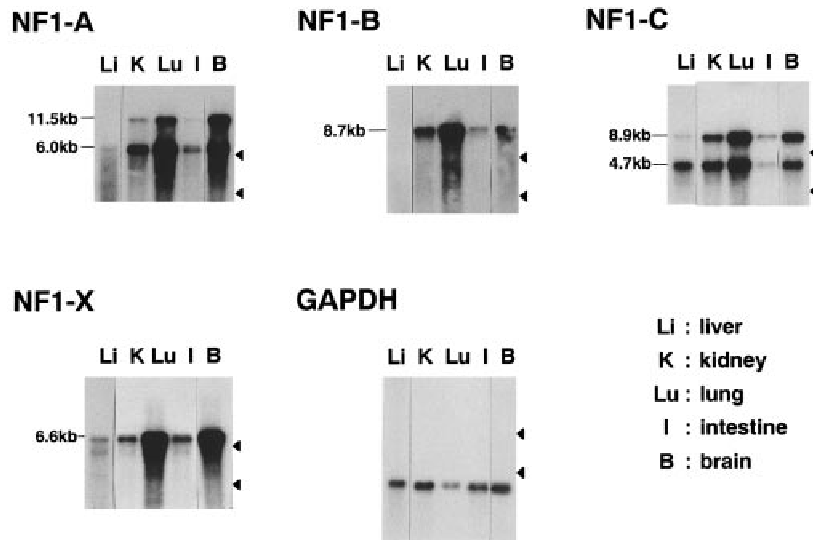

Figure 2 Expression profiles of NF1 mRNA species by Northern blot analysis

Poly(A)$^+$ RNA (2 μg) from various rat tissues was loaded in each lane. The filter was hybridized with the radiolabelled NF1-specific probes. The same filter was subsequently hybridized with a labelled glyceraldehyde-3-phosphate dehydrogenase (GAPDH) probe. Arrowheads indicate the positions of 28 S and 18 S rRNA.

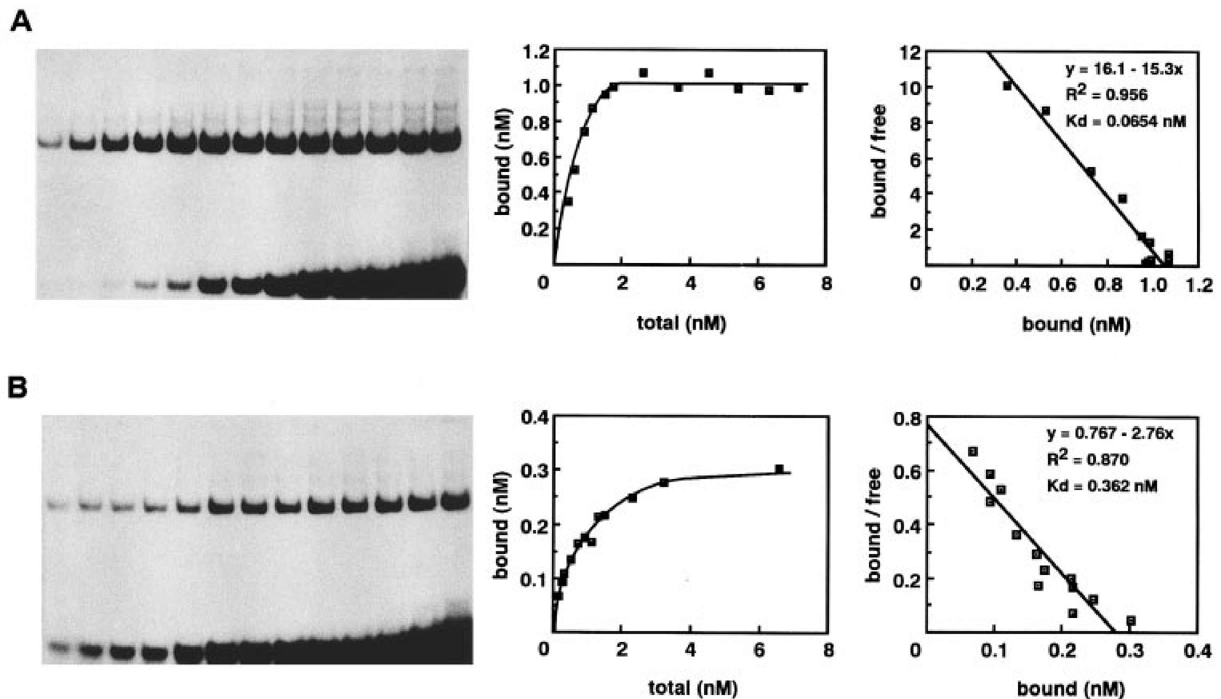

Figure 3 Determination of the dissociation constants of NF1-A proteins

The left panels show the results of the gel-shift analyses with increasing amounts of NF1 sites of the adenovirus replication origin (**A**) and the GST-P gene (**B**). The amount of bound probe was plotted against the total input (middle panels). A Scatchard analysis was performed by plotting the ratio of bound to free protein against bound protein (right panels).

random site selection [27]; in other words, the adenovirus replication origin showed the highest affinity. This result demonstrated that the specificities of the NF1 proteins are indistinguishable. Surprisingly, however, the affinities of NF1-A and NF1-X were higher than those of NF1-B and NF1-C. It was confirmed that GST-NF1 fusion proteins are equally stable in this experiment as follows. The protein fractions were left for 30 min at room temperature, under the same conditions as for the binding reaction for the *in vitro* DNA binding studies; these fractions were then subjected to electrophoresis and stability was checked by both Coomassie Brilliant Blue staining and immunoblot analysis in which the anti-NF1 antibody from Santa Cruz

© 1999 Biochemical Society

Table 1 Dissociation constants for NF1 proteins for the four NF1 sites

Results are means ± S.D. for three to nine independent experiments.

	Dissociation constant (pM)			
Probe	NF1-A	NF1-B	NF1-C	NF1-X
Adenovirus replication origin	56 ± 7	160 ± 47	170 ± 9	140 ± 10
α-Globin	210 ± 36	480 ± 150	420 ± 120	290 ± 13
HMG-CoA reductase	280 ± 28	510 ± 94	610 ± 150	220 ± 6
GST-P	280 ± 79	370 ± 110	610 ± 84	180 ± 42

was used. In neither analysis was the degradation band observed (results not shown).

Effect of the C-terminal regions of NF1-B proteins on various promoter activities

Previous work has shown that NF1-A contains a repression domain enriched with proline and serine residues and that NF1-C and NF1-X activate the transcription mediated by the proline-rich transcriptional activation domains [10,11,13]. We examined the effect of the C-terminal regions of NF1-B isoforms on the four kinds of gene promoter activity (Figure 4). Because NF1 proteins were expressed in almost all culture cells, we fused the C-terminus of NF1-B proteins to the DNA-binding domain of GAL4 (1–147) to exclude the effect of endogenous NF1 proteins on the transcription. GAL4–A1(145–509) was used, because the region 145–509 of NF1-A1 showed the most effective repression activity [13]. Three isoforms of the human NF1-C cDNA species, NF1-C1, NF1-C2 and NF1-C3, have been cloned [10]. The transcriptional regulatory domain of only NF1-C1 has been characterized well, and the proline-rich transcriptional activation domain that was the C-terminal end was identified. We cloned rat NF1-C2, which is similar to human NF1-C2; rat NF1-C1 has not been isolated. Therefore we examined the effect of the C-terminal end of rat NF1-C2 as well as human NF1-C1 on the promoter activities. The expression vectors were co-transfected into HeLa cells with a reporter plasmid containing the GST-P promoter and five binding sites for the GAL4 DNA-binding domain (Figure 4B). The results were shown as luciferase activity relative to that of the GAL4 DNA-binding domain alone. The GST-P promoter has a PMA response element and a GC-box that are binding sites for AP1 and Sp1 respectively [30]. The C-terminal half of NF1-B2 and NF1-B3, but not NF1-B1, slightly activated the transcription activity of the GST-P gene. The C-terminal regions of NF1-A1 decreased the transcription activity when co-transfected with GAL4-A1 as an expression vector. The NF1-C isoforms hardly activated the GST-P promoter activity. Similar results were obtained when the reporter plasmids containing SV40 and C/EBPδ promoters were used (Figures 4C and 4D). However, the C-terminal regions of all of the NF1 isoforms repressed the promoter activity of the MTIIA gene (Figure 4E). These observations indicate that the transcriptional regulatory function of the NF1 isoforms is affected by the factors that bind to the promoter regions.

To show that the functional differences in the NF1 GAL4 fusion proteins were not due to differences in their relative expression levels, we performed a Western blot analysis with the anti-GAL4 antibody. All samples were standardized by the β-galactosidase activity, except GAL4-A1(145–509) (55.7 kDa) and GAL4-B1(239–561) (52.1 kDa), which were respectively 5-fold

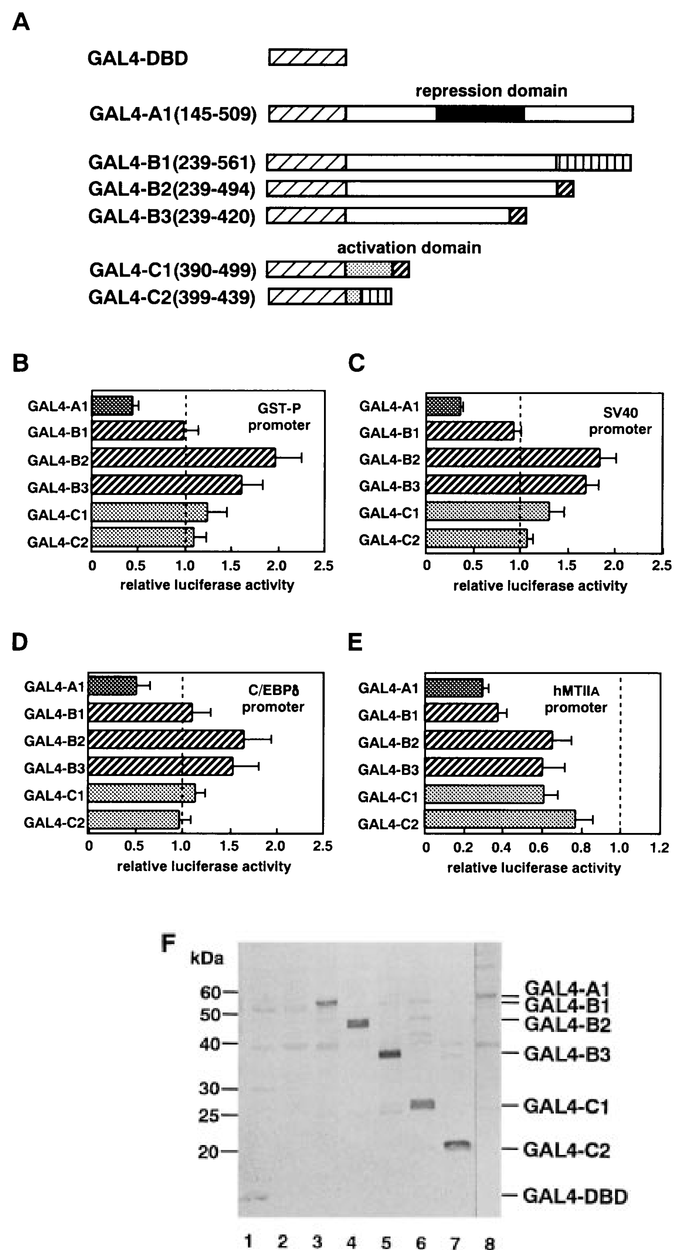

Figure 4 Effect of the C-terminal regions of NF1 isoforms on the transcription activity of several promoters

(**A**) Diagrams of the effector plasmids containing the GAL4 DNA-binding domain and the C-terminal regions of various NF1 isoforms. The start and end points of residues of NF1 proteins inserted are indicated in parentheses in the construct name. The transcriptional repression domain in NF1-A1 is highlighted by a filled box. The NF1-B1- and NF1-C2-specific sequences are marked by vertical hatching in the respective constructs. The diagonally hatched boxes indicate the Ser-Trp-Tyr-Leu-Gly motifs. The transcriptional activation sequences are marked by stippled boxes. All of the cDNA species are derived from the rat except human NF1-C1. Each dish received the effector plasmid and β-galactosidase expression vector as well as the reporter plasmid containing GST-P (**B**), SV40 (**C**), C/EBPδ (**D**) or MTIIA (**E**) promoter, and GAL4 binding sites. The amino acid positions of the effector plasmids were omitted (**B**–**E**). Luciferase activity was measured in the cell extracts and normalized by the β-galactosidase activity. The results are indicated by the luciferase activity (means and S.D.; $n = 3$–8) relative to the activity produced by co-transfection of the GAL4 DNA-binding domain alone. (**F**) Expression levels of GAL4–NF1 fusion proteins. The nuclear extracts were resolved by SDS/PAGE and immunoblotted with anti-(GAL4-DBD) antibody, where DBD is the DNA-binding domain. All samples are standardized by the β-galactosidase activity, except GAL4-A1(145–509) and GAL4-B1(239–561), which respectively contained 5-fold and 6-fold the amount of the others. Lane 1, GAL4-DBD; lane 2, GAL4-A1(145–509); lane 3, GAL4-B1(239–561); lane 4, GAL4-B2(239–494); lane 5, GAL4-B3(239–420); lane 6, GAL4-C1(390–499); lane 7, GAL4-C2(399–439); lane 8, overstained for GAL4-A1(145–509).

© 1999 Biochemical Society

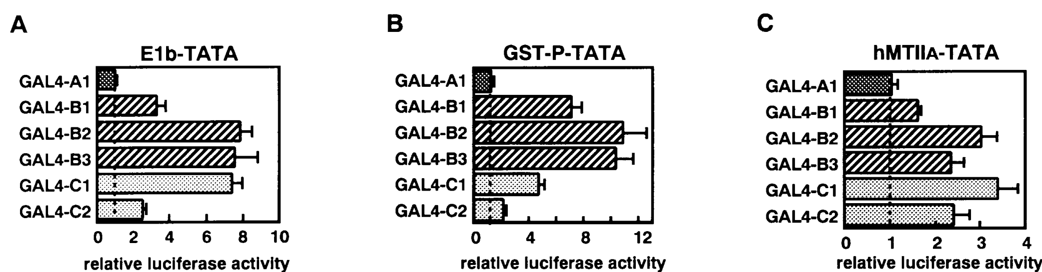

Figure 5 Effect of the C-terminal regions of NF1 isoforms on the transcription activity of minimal promoters

The effector plasmids containing the GAL4 DNA-binding domain and the C-terminal region of various NF1s were transfected with E1b (**A**), GST-P (**B**) or MTIIA (**C**) minimal promoter. Relative luciferase activities (means and S.D.) are shown as in Figure 4.

and 6-fold the amount of the others. The expression levels of GAL4-B2(239–494), GAL4-B3(239–420), GAL4-C1(390–499) and GAL4-C2(399–439) were similar, as shown in Figure 4(F). The expression levels of GAL4-A1(145–509) and GAL4-B1(239–561) seem to be lower than those of others, although the transfer efficiency of the protein above 50 kDa to nitrocellulose membrane was lower than that of proteins below 50 kDa. It is therefore likely that GAL4-A1(145–509) and GAL4-B1(239–561) would have had a stronger repression activity in Figure 4 if the fusion proteins had been expressed at the same levels as the others.

Activation of minimal promoter activities by the C-terminal half of NF1-B proteins

To define the function of the C-terminal regions of the NF1-B isoforms, a luciferase reporter gene, driven by minimal promoters consisting of the adenovirus E1b, GST-P or MTIIA gene TATA box and five binding sites for the GAL4 DNA-binding domain, was used because the transcriptional regulatory function of the NF1 isoforms is affected by the factors that bind to the promoter regions. GAL4-C1(390–499), containing the proline-rich transcriptional activation domain of human NF1-C, activated the minimal promoter activity (Figure 5). However, GAL4-A1(145–509), containing the GAL4 DNA-binding domain and the C-terminus of NF1-A1, hardly repressed the minimal promoter activities. The transcription activities increased when the reporter plasmid was co-transfected with GAL4 fused to the NF1-B isoforms as expression vectors. The C-terminal half of NF1-B2 and NF1-B3 activated the transcription more effectively than did that of NF1-B1. These effects of the NF1 members on the promoter activity were not dependent on the type of minimal promoter used. Therefore the C-terminus of NF1-B can function as an activator of transcription in a DNA-binding-site-dependent manner.

Identification of the transcriptional activation domain of NF1-B

To define the domain that is important for activation, we constructed a series of deletion mutants of the C-terminal regions of NF1-B (Figure 6). Deletion of the regions 239–308 and 239–391 did not affect the activation activity of NF1-B1, because the relative luciferase activities of GAL4-B1(309–561) and GAL4-B1(392–561) were the same as that of GAL4-B1(239–561) (see Figure 6A, lanes 2–4). The activation activity was abolished when the remaining region 392–561 was separated into two regions [see GAL4-B1(392–489) and GAL4-B1(490–561); Figure 6A, lanes 5 and 6]. These observations indicate that the transactivation domain of NF1-B1 exists in the region 392–561. Western blot analysis with anti-(GAL4 DNA-binding domain) antibody revealed that similar levels of GAL4 NF1-B fusion proteins were expressed in the transfected HeLa cells (Figure 6B). GAL4-B1(309–561) might show an effective transactivating activity, because the relative expression level of GAL4-B1(309–561) was lower than those of others.

We also tested the activity of the GAL4 DNA-binding domain fused to several portions of NF1-B2 and NF1-B3. The deletion of the region 239–308 did not affect the transactivation activities of GAL4-B2(239–494) and GAL4-B3(239–420) [see GAL4-B2(309–494) and GAL4-B3(309–420); Figure 6A, lanes 7, 8, 10 and 11]. The transactivation activities of NF1-B2 and NF1-B3 were decreased by deleting the region 239–391 [see GAL4-B2(392–494) and GAL4-B3(392–420); Figure 6A, lanes 9 and 16]. These results mean that the region 309–391 is important for the transactivation of NF1-B2 and NF1-B3. However, the region 309–391 did not itself show the same activation activity as that seen in the overall region of C-terminus of NF1-B2 and NF1-B3 [see GAL4-B1/2/3(309–391); Figure 6A, lane 18]. When this region was divided into two regions, 309–354 and 355–391, these fragments did not show the activation activities, either (see Figure 6A, lanes 19 and 20).

The short motif Ser-Trp-Tyr-Leu-Gly at the C-terminus of NF1-B3 is conserved in NF1-A1, NF1-B2 and NF1-C1 ([3,7]; see also Figure 1). GAL4-B1/2/3(309–415), which does not contain the Ser-Trp-Tyr-Leu-Gly motif, was without activation activity (see Figure 6A, lane 17), whereas GAL4-B3(309–420), which includes the Ser-Trp-Tyr-Leu-Gly motif, has high activity (see Figure 6A, lane 11). Therefore the Ser-Trp-Tyr-Leu-Gly motif is essential to the transactivation. However, this motif itself is not a sufficient minimum activation domain, because the C-terminal 29 residues, GAL4-B3(392–420), did not activate the luciferase activity (see Figure 6A, lane 16).

To define more narrowly the NF1-B3 activation domain, we prepared several deletion constructs. The transactivation activity of NF1-B3 was decreased with progressive deletions, GAL4-B3(325–420), GAL4-B3(340–420), GAL4-B3(355–420) and GAL4-B3(373–420) (see Figure 6A, lanes 12–15). We conclude that transactivations by NF1-B2 and NF1-B3 are required for the region 309 to the stop codon. These observations indicate that this region is essential but not sufficient for the activation mediated by NF1-B2 and NF1-B3, although the region 309–391 is not required for transactivation by NF1-B1.

© 1999 Biochemical Society

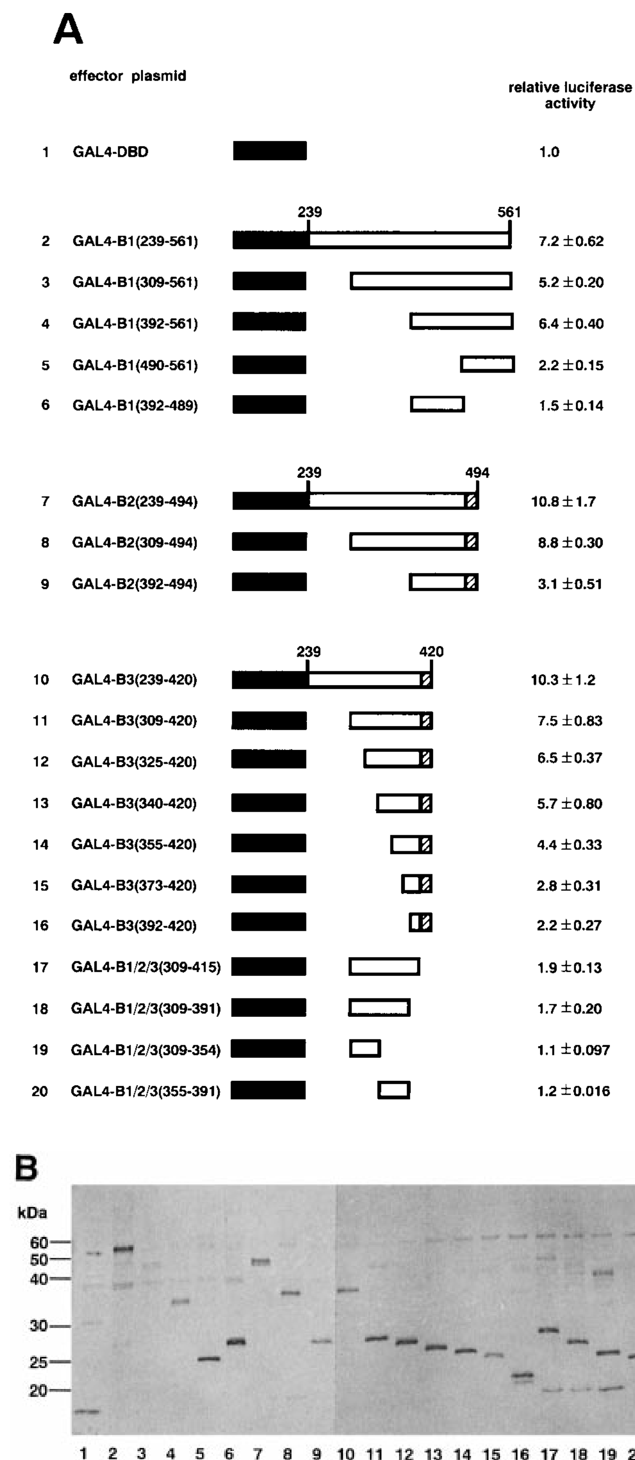

Figure 6 Identification of the NF1-B activation domain

(**A**) At the left are shown the effector plasmids that contain the GAL4 DNA-binding domain and the C-terminal region of various NF1-Bs. These constructs are termed GAL4-B(*x*–*y*), where *x* and *y* indicate the start and end residues respectively of the inserted NF1-B proteins. GAL4-B1(239–561), GAL4-B2(239–494) and GAL4-B3(239–420) contain the intact region from residue 239 to the stop codon of each NF1-B. Cells were co-transfected with the effector plasmids together with the reporter plasmid containing GST-P minimal promoter and GAL4 DNA-binding sites. Relative luciferase activities (means ± S.D.) are shown at the right. The Ser-Trp-Tyr-Leu-Gly motifs in NF1-B2 and NF1-B3 are indicated by the hatched boxes. (**B**) Expression levels of GAL4–NF1 fusion proteins. Western blot analysis with antibody against the GAL4 DNA-binding domain was performed for the GAL4 fusion proteins in transfected HeLa cells. All samples are normalized by the β-galactosidase activity, except GAL4-B1(239–561) (lane 2), which contained 6-fold the amount of the others. Lane numbers correspond to the numbering of GAL4 fusion proteins as in (**A**).

DISCUSSION

More than 20 splicing isoforms of NF1 have been cloned from several species. In the present study we isolated five splicing isoforms derived from rat NF1-B, NF1-C and NF1-X genes. These clones have quite high similarities to the equivalent parts of human, mouse and hamster NF1s, sharing more than 97 % identity in amino acid sequences [3,5,31]. NF1-B1, NF1-B3, NF1-C2 and NF1-X1 are the rat counterparts of hamster NF1-B, human NF1-B2, human NF1-C2 and hamster NF1-X respectively [3,5,32]. NF1-B2 is a novel type of NF1 gene product. Two and three types of NF1-B isoform were cloned from chicken and human respectively; the N-terminal regions of them are conserved [7,32]. Qian et al. [8] reported the only human NF1-B gene in the human genone, and its chromosomal location was mapped to position 9p24.1. We speculate that the N-terminal regions of the NF1-B isoforms in rat are also conserved, although we have not isolated the clones that contain the full-length NF1-B cDNA species.

Although many NF1 cDNA species were cloned, the expression profile of NF1 is unclear. RNA was isolated from rat tissues and analysed for the expression of all NF1 genes. With NF1-B and NF1-X, one major signal was detected. Two transcripts, 11.5 and 6.0 kb of NF1-A mRNA and 8.9 and 4.7 kb of NF1-C, were expressed. NF1 mRNA species were detected ubiquitously in various rat tissues, although the level of NF1 was higher in the lung and lower in the liver. Chaudhry et al. [31] recently reported that in adult mice, all four NF1 genes are expressed most highly in the lung, liver, heart, and other tissues but only weakly in the spleen and testes. Although the sizes, numbers and expression profiles of rat NF1 mRNA species resembled those in mice, the level of mRNA in the liver was different. The reason for this difference is unknown.

NF1 family proteins bind to the specific sequences and are required for tissue- or cell-type-specific expression. Although all four NF1 proteins bind to the adenovirus replication origin, the specificities and activities of NF1 proteins at the NF1 sites in the promoter regions are unclear. We measured the dissociation constants for each protein for the four NF1 sites including the adenovirus, GST-P, α-globin and HMG-CoA reductase genes. The order of affinity for the four NF1 sites in all four of the NF1 proteins depended on the similarity to the NF1-A consensus sequence derived from the PCR-mediated random site selection (i.e. the adenovirus replication origin showed the highest affinity). This observation means that the specificities of the NF1 proteins are indistinguishable. Surprisingly, the binding affinities of NF1-A and NF1-X were higher than those of NF1-B and NF1-C. We have showed previously that the perfect palindrome sequence containing TTGGCA was the most effective competitor in the NF1 sites [27]. Although the perfect palindrome sequence probably exists in the genomic DNA, such a sequence has not been found in the functional region in the genes. The highest binding activity might not be required for the regulation of replication and transcription, or rather, a stringent binding specificity might be inconvenient. Moreover, the regulation of transcription is not required for the high affinity that was detected in the adenovirus replication origin. This means that both the release of the factor from DNA and its binding to DNA are important in the regulation of transcriptional functions. The affinities of NF1-B and NF1-C were lower than those of the other NF1 proteins, indicating that NF1-B and NF1-C are more easily competed for by other NF1 proteins. The difference in DNA-binding activities might contribute to the regulation of the gene expression by NF1 family proteins.

The transcriptional regulatory regions of rat NF1-A, human

© 1999 Biochemical Society

NF1-C and *Xenopus* NF1-X have been examined by using heterologous DNA-binding domains [10,11,13]. The function of the C-terminal region of NF1-B is unknown. Although the activities of the minimal promoters were increased by NF1-B2 and NF1-B3, the C-terminal regions of NF1-B2 and NF1-B3 slightly increased the promoter activities of the GST-P, SV40 and C/EBPδ genes and decreased MTIIA promoter activity. These promoters have several *cis*-elements, including a PMA response element, a GC box and an AP2 site [30,33–35]. NF1-B1, NF1-C1 and NF1-C2 were able to activate the minimal promoter activities but did not induce the activities derived from the promoters containing several *cis*-elements. NF1-A repressed all promoter activities and did not activate the minimal promoter activities. These observations indicate that the function of transcriptional regulation by NF1 proteins is influenced by the factors that bind to the *cis*-elements near the NF1 sites. Indeed, NF1 sites are located within many transcription factor sites [9,36–40]. NF1 proteins supplement or repress the transactivation mediated by transcriptional activators such as AP1 and Sp1, rather than activating the transcriptional activity constantly.

The repression activity of NF1-A is higher than those of NF1-B and NF1-C. We have previously purified NF1-A and NF1-B from rat liver as factors that bind to the silencer in the GST-P gene [9]. The promoter activity of the GST-P gene was increased by NF1-B2 and NF1-B3, but not NF1-B1, and decreased by NF1-A1. The balance of the expression of the NF1 isoforms might be important for the silencer activity of the GST-P gene. C/EBPα, one of the GST-P-silencer-binding proteins, represses GST-P promoter activity and attenuates the transcriptional stimulation by C/EBPβ. It is likely that the ratio of C/EBPα to C/EBPβ is one of the important factors in GST-P silencer activity. Indeed, a decrease in this ratio resulted in an increase in GST-P expression during hepatocarcinogenesis [22]. In the near future, it may be possible to use the specific antibodies against the specific peptide sequences in each NF1 isoform to dissolve the expression profiles of NF1 family proteins in the liver during hepatocarcinogenesis.

We have previously identified a rat NF1-A repression region that is divided into two 100-residue domains, one enriched in serine and glycine residues and the other in proline and serine residues [13]. The proline-rich activation domain was identified in the C-terminal region of human NF1-C and *Xenopus* NF1-X [10,11]. To define the function of the C-terminal regions of NF1-B isoforms, a luciferase reporter gene driven by minimal promoters was used. In the present study we found the activation domains of NF1-B isoforms. Transactivations by NF1-B2 and NF1-B3 are required for the region from residue 309 to the stop codon. However, the region 309–391 is not essential for the activation by NF1-B1 that requires the region 392 to the stop codon. The C-termini of NF1 proteins are diverse but the transcriptional regulatory regions are similar in their amino acid components. The activation domains of NF1-B isoforms have abundant proline and serine residues. The region 309–391, in particular, is rich in proline (25%) and serine (17%) residues.

The transactivation activity of NF1-B3 was gradually decreased with progressive deletions from 309 to 391, indicating that all of the C-terminal 112 residues of NF1-B3 are needed as the activation domain. On the activation domain of the N-terminal 100 residues of the transcription factor Yin and Yang 1 (YY1), the activation activity was abolished by a series of progressive deletion mutants [41]. It is thought that all of the regions and the normal structure of these activation domains might be important for maximum activation. The short motif Ser-Trp-Tyr-Leu-Gly at the C-terminus of NF1-B3 was also required for activation. The Trp-Arg-Pro-Trp motif of the Hairy-related basic helix–loop–helix is sufficient to interact with Groucho or TLE mammalian counterparts [42]. The Ser-Trp-Tyr-Leu-Gly motif conserved in NF1-A1, NF1-B2, NF1-B3 and NF1-C1 might be important for protein–protein interaction.

NF1-C2 showed the same degrees of activation activity on the E1b and GST-P minimal promoters. In contrast, the NF1-C1 and NF1-B isoforms increased more effectively the minimal promoter activity of the E1b and GST-P genes respectively. Chaudhry et al. [43] reported that NF1-X shows the strongest activation of the adenovirus major late promoter and NF1-B is the most potent activator of the mouse mammary tumour virus long terminal promoter. These promoter specificities by NF1 isoforms might be derived from the targets of the activation domains of NF1 proteins. The proline-rich activation domain of human NF1-C1 selectively interacts with TFIIB, TBP, TBP-associated factor 55 and histone H3 [14–17]. To define the difference between transcriptional regulations by NF1 members, we are attempting to identify the factors interacting with the transcriptional regulatory regions of NF1 family members.

We thank Dr. N. Miura at Akita University (Akita, Japan) for kindly providing the rat liver λgt10 cDNA library. This work was supported in part by grants from the Ministry of Education, Science, Sports and Culture, Japan, and from the Nissan Science Foundation.

REFERENCES

1 Wedel, A. and Ziegler-Heitbrock, H. W. (1995) Immunobiology **193**, 171–185
2 Mitchell, P. J. and Tjian, R. (1989) Science **245**, 371–378
3 Santoro, C., Mermod, N., Andrews, P. C. and Tjian, R. (1988) Nature (London) **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. Natl. Acad. Sci. U.S.A. **85**, 8963–8967
6 Inoue, T., Tamura, T., Furuichi, T. and Mikoshiba, K. (1990) J. Biol. Chem. **265**, 19065–19070
7 Kruse, U. and Sippel, A. E. (1994) J. Mol. Biol. **238**, 860–865
8 Qian, F., Kruse, U., Lichter, P. and Sippel, A. E. (1995) Genomics **28**, 66–73
9 Osada, S., Daimon, S., Ikeda, T., Nishihara, T., Yano, K., Yamasaki, M. and Imagawa, M. (1997) J. Biochem. (Tokyo) **121**, 356–363
10 Mermod, N., O'Neill, E. A., Kelly, T. J. and Tjian, R. (1989) Cell **58**, 741–753
11 Roulet, E., Armentero, M. T., Krey, G., Corthesy, B., Dreyer, C., Mermod, N. and Wahli, W. (1995) Mol. Cell. Biol. **15**, 5552–5562
12 Dekker, J., van Oosterhout, J. A. and van der Vliet, P. C. (1996) Mol. Cell. Biol. **16**, 4073–4080
13 Osada, S., Ikeda, T., Xu, M., Nishihara, T. and Imagawa, M. (1997) Biochem. Biophys. Res. Commun. **238**, 744–747
14 Kim, T. K. and Roeder, R. G. (1994) Proc. Natl. Acad. Sci. U.S.A. **91**, 4170–4174
15 Xiao, H., Lis, J. T., Xiao, H., Greenblatt, J. and Friesen, J. D. (1994) Nucleic Acids Res. **22**, 1966–1973
16 Chiang, C. M. and Roeder, R. G. (1995) Science **267**, 531–536
17 Alevizopoulos, A., Dusserre, Y., Tsai-Pflugfelder, M., von der Weid, T., Wahli, W. and Mermod, N. (1995) Genes Dev. **9**, 3051–3066
18 Hattori, M. and Sakaki, Y. (1986) Anal. Biochem. **152**, 232–238
19 Sadowski, I. and Ptashne, M. (1989) Nucleic Acids Res. **17**, 7539
20 Chen, C. and Okayama, H. (1987) Mol. Cell. Biol. **7**, 2745–2752
21 Herbomel, P., Bourachot, B. and Yaniv, M. (1984) Cell **39**, 653–662
22 Osada, S., Takano, K., Nishihara, T., Suzuki, T., Muramatsu, M. and Imagawa, M. (1995) J. Biol. Chem. **270**, 31288–31293
23 Novak, A., Goyal, N. and Gronostajski, R. M. (1992) J. Biol. Chem. **267**, 12986–12990
24 Meisterernst, M., Rogge, L., Foeckler, R., Karaghiosoff, M. and Winnacker, E. L. (1989) Biochemistry **28**, 8191–8200
25 Kruse, U. and Sippel, A. E. (1994) FEBS Lett. **348**, 46–50
26 Rupp, R. A. W. and Sippel, A. E. (1987) Nucleic Acids Res. **15**, 9707–9726
27 Osada, S., Daimon, S., Nishihara, T. and Imagawa, M. (1996) FEBS Lett. **390**, 44–46
28 Xu, M., Osada, S., Imagawa, M. and Nishihara, T. (1997) J. Biochem. (Tokyo) **122**, 795–801
29 Puzianowska-Kuznicka, M. and Shi, Y. B. (1996) J. Biol. Chem. **271**, 6273–6282
30 Sakai, M., Okuda, A. and Muramatsu, M. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 9456–9460

© 1999 Biochemical Society

31 Chaudhry, A. Z., Lyons, G. E. and Gronostajski, R. M. (1997) Dev. Dyn. **208**, 313–325
32 Liu, Y., Bernard, H. U. and Apt, D. (1997) J. Biol. Chem. **272**, 10739–10745
33 Lee, W., Haslinger, A., Karin, M. and Tjian, R. (1987) Nature (London) **325**, 368–372
34 Imagawa, M., Chiu, R. and Karin, M. (1987) Cell **51**, 251–260
35 Yamada, T., Tobita, K., Osada, S., Nishihara, T. and Imagawa, M. (1997) J. Biochem. (Tokyo) **121**, 731–738
36 Xanthopoulos, K. G., Prezioso, V. R., Chen, W. S., Sladek, F. M., Cortese, R. and Darnell, Jr., J. (1991) Proc. Natl. Acad. Sci. U.S.A. **88**, 3807–3811
37 Short, M. K., Clouthier, D. E., Schaefer, I. M., Hammer, R. E., Magnuson, M. A. and Beale, E. G. (1992) Mol. Cell. Biol. **12**, 1007–1020
38 Garlatti, M., Tchesnokov, V., Daheshia, M., Feilleux-Duche, S., Hanoune, J., Aggerbeck, M. and Barouki, R. (1993) J. Biol. Chem. **268**, 6567–6574
39 Bernier, D., Thomassin, H., Allard, D., Guertin, M., Hamel, D., Blaquiere, M., Beauchemin, M., LaRue, H., Estable-Puig, M. and Belanger, L. (1993) Mol. Cell. Biol. **13**, 1619–1633
40 Spitz, F., Salminen, M., Demignon, J., Kahn, A., Daegelen, D. and Maire, P. (1997) Mol. Cell. Biol. **17**, 656–666
41 Bushmeyer, S., Park, K. and Atchison, M. L. (1995) J. Biol. Chem. **270**, 30213–30220
42 Fisher, A. L., Ohsako, S. and Caudy, M. (1996) Mol. Cell. Biol. **16**, 2670–2677
43 Chaudhry, A. Z., Vitullo, A. D. and Gronostajski, R. M. (1998) J. Biol. Chem. **273**, 18538–18546

Received 17 February 1999/12 May 1999; accepted 8 June 1999

© 1999 Biochemical Society