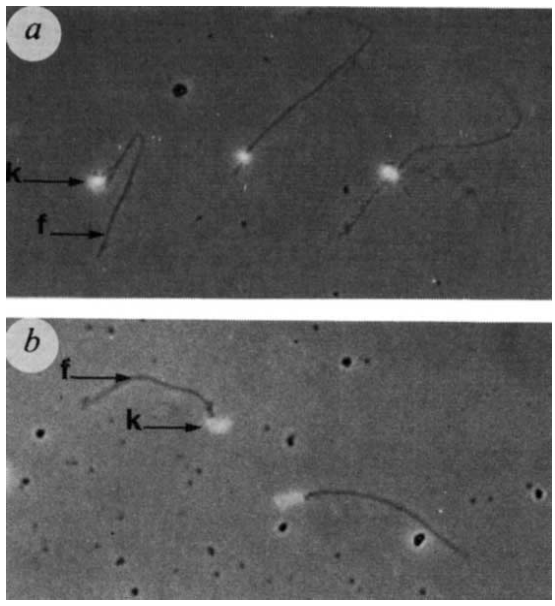


FIG. 4 Combined phase-contrast and DAPI fluorescence micrographs of isolated *T. brucei* flagella show kinetoplasts (bright fluorescent dots) still attached to basal bodies. *a*, Flagella isolated by hypotonic lysis and cytoskeletal depolymerization. *b*, Flagella isolated by detergent lysis and cytoskeletal depolymerization. The isolated flagella (f) and DAPI stained kinetoplasts (k) are viewed by combined phase-contrast/fluorescence microscopy.
METHODS. Hypotonic lysis: log-phase cell cultures were collected by centrifugation, washed three times in 0.25 M sucrose, 20 mM Tris–HCl, pH 7.9. Microtubule cytoskeletal depolymerization and cell lysis was achieved by resuspending and incubating cells in 3 mM Ca^{2+} in double distilled water for 3 h at 4 °C. Detergent lysis: EDTA was added to a mid-log phase culture, $4–5 \times 10^6$ cells ml^{-1}, final concentration of 5 mM. Cells were then collected by centrifugation, washed in PBS, resuspended on ice for 10 min in extraction buffer, (0.5% Triton X-100 in PMN (10 mM NaH_2PO_4, 150 mM NaCl, 1 mM $MgCl_2$ pH7.2)). Cytoskeletons were collected by centrifugation, washed in extraction buffer, and resuspended on ice for 45 min in 1 mM Ca^{2+} in PMN. Isolated flagella are DAPI-stained and viewed by fluorescence microscopy.

Our experiments reveal an established structural and functional link between the mitochondrial genome and basal bodies in the *T. brucei* cell. This link provides a means of ensuring high fidelity segregation of these single-copy cytoplasmic organelles to daughter cells. By contrast, other eukaryotic cells possess multiple and dispersed mitochondrial genomes which may not require such stringent segregation fidelity. The high-order system present in *T. brucei* may represent a remnant of an evolutionarily archaic mechanism for DNA segregation. □

Received 5 June; accepted 27 June 1991.

1. Vickerman, K. *Trans. R. Soc. trop. Med. Hyg.* **56,** 487–495 (1962).
2. Robertson, M. *Phil. Trans. R. Soc.* B**203,** 161–184 (1913).
3. Simpson, L. *J. Protozool.* **15,** 132–136 (1968).
4. Molyneux, D. H. *Parasitology* **59,** 55–66 (1969).
5. Lewis, D. H. *J. Protozool.* **22,** 344–352 (1975).
6. Steinert, M., Van Assel, S. & Steinert, G. in *Biochemistry of Parasites and Host-Parasite Relationships* (ed. Van den Bossche, H.) 193–202 (Elsevier/North-Holland Biomedical, Amsterdam, 1976).
7. Souto-Padron, T., De Souza, W. & Heuser, J. E. *J. Cell Sci.* **69,** 167–178 (1984).
8. Hiruki, T. *Zentbl. Bakt. Mikrobiol. Hyg.* **264,** 392–398 (1987).
9. Woods, A. *et al. J. Cell Sci.* **93,** 491–500 (1989).
10. Sherwin, T. & Gull, K. *Phil. Trans. R. Soc.* B**323,** 573–588 (1989).
11. Chow, K. C., Macdonald, T. L. & Ross, W. E. *Molec. Pharmacol.* **34,** 467–473 (1988).
12. Patel, S., Austin, C. A. & Fisher, L. M. *Anticancer Drug Des.* **5,** 149–157 (1990).
13. Shapiro, T. A. & Englund, P. T. *Proc. natn. Acad. Sci. U.S.A.* **87,** 950–954 (1990).
14. Ootso, K. *et al. Cancer Res.* **40,** 1707–1717 (1980).
15. Brun, R. & Schonenberger, M. *Acta trop.* **36,** 289–292 (1979).
16. Sherwin, T., Schneider, A., Sasse, R., Seebeck, T. & Gull, K. *J. Cell Biol.* **104,** 439–446 (1987).

ACKNOWLEDGEMENTS. We thank L. M. Fisher for teniposide. Ansamitocin was a donation from Takeda Chemical Industries of Japan. D.R. is supported by the SERC and this investigation received support from the UNDP/World Bank/WHO special programme for Research and Training in Tropical Diseases.

Cloning of an NF-κB subunit which stimulates HIV transcription in synergy with p65

Roland M. Schmid, Neil D. Perkins, Colin S. Duckett, Philip C. Andrews & Gary J. Nabel*

Howard Hughes Medical Institute, Departments of Internal Medicine and Biological Chemistry, University of Michigan Medical Center, Ann Arbor, Michigan 48109-0650, USA
* To whom correspondence should be addressed

THE transcription factor NF-κB is a protein complex which comprises a DNA-binding subunit and an associated transactivation protein (of relative molecular masses 50,000 (50K) and 65K, respectively)[1,2]. Both the 50K and 65K subunits have similarity with the *rel* oncogene and the *Drosophila* maternal effect gene *dorsal*[3–6]. The 50K DNA-binding subunit was previously thought to be a unique protein, derived from the 105K gene product (p105). We now report the isolation of a complementary DNA that encodes an alternative DNA-binding subunit of NF-κB. It is more similar to p105 NF-κB than other family members and defines a new subset of *rel*-related genes. It is synthesized as a ~100K protein (p100) that is expressed in different cell types, contains cell cycle motifs and, like p105, must be processed to generate a 50K form. A 49K product (p49) can be generated independently from an alternatively spliced transcript; it has specific κB DNA-binding activity and can form heterodimers with other *rel* proteins. In contrast to the ~50K protein derived from p105, p49 acts in synergy with p65 to stimulate the human immunodeficiency virus (HIV) enhancer in transiently transfected Jurkat cells. p49/p100 NF-κB could therefore be important in the regulation of HIV and other κB-containing genes.

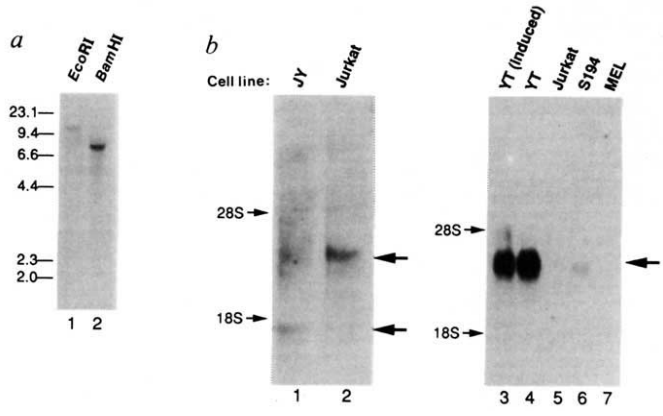
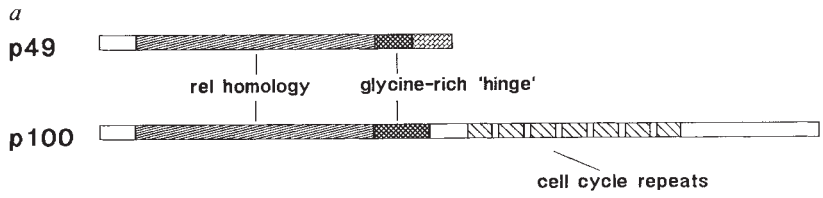


FIG. 1 DNA and RNA analysis of the p49 cDNA clone. *a*, Southern blot analysis of p49 on human genomic DNA (10 μg) cut with *Eco*RI (lane 1) or *Bam*HI (lane 2). The DNA probe used was a 1,069-base-pair *Bst*XI fragment of p49 which lacks both 5′-untranslated DNA and the repetitive region from the 3′ end of the clone. Molecular size markers (in kb) are indicated on the left. *b*, Northern blot analysis of 10 μg poly(A)$^+$ RNA from JY, an Epstein–Barr virus-transformed B-cell line (lane 1) (from J. Leiden), and the Jurkat T leukaemia line (lane 2) using the *Bst*XI fragment of p49 as a probe with GeneScreen Plus (DuPont). Another analysis was performed using 10 μg poly(A)$^+$ RNA from TPA-stimulated and unstimulated YT T leukaemia cells (lanes 3 and 4); the Jurkat T leukaemia line (lane 5); S194, a murine B cell line (lane 6); and the murine erythroleukaemia (MEL), cell line (lane 7), on a nitrocellulose filter. Migration positions of ribosomal RNA markers are indicated (28S and 18S). Arrows denote specific hybridizable mRNA species of 1.9 kb (lanes 1, 2) and ~3.5 kb (lanes 1–7). For each blot, comparable amounts of actin RNA were detected in all lanes using a human β-actin probe.

We have isolated cDNAs encoding the NF-κB subunit proteins in order to characterize them and determine the mechanism of activation by NF-κB. Trypsin-digested κB binding proteins purified from bovine spleen were sequenced. Three peptides were identical to a DNA-binding subunit of human NF-κB/KBF1 (ref. 3); another was 75% identical to a homologous peptide of p105. This peptide sequence suggested that an alternative NF-κB protein could exist.

A polymerase chain reaction (PCR) probe amplified from degenerate primers was found to be identical to a 700-bp fragment of human c-*rel* (refs 7, 8), and was used for low-stringency hybridization. From 6×10^5 recombinants, two clones were identified that hybridized at low stringency. By Southern blot analysis at high stringency, we found that this gene was present in single copy (Fig. 1*a*), and northern blotting revealed at least two hybridizable RNA species of ~1.9 kilobases (kb) and ~3.5 kb

a



b

```
p49  MESCYNPGLDGIIEYDD---------FKLNSS---------IVEPKEPAPETADGPYLVI
p100 MESCYNPGLDGIIEYDD---------FKLNSS---------IVEPKEPAPETADGPYLVI
p105 MAE------------DDPYLGRPEQMFHLDPSLTHTIFNPEVFQPQMALP-TADGPYLQI
     M              DD         F L  S           P    P TADGPYL I

p49  VEQPKQRGFRFRYGCEGPSHGGLPGASSEKGRKTYPTVKICNYEGPAKIEVDLVTHSDPP
p100 VEQPKQRGFRFRYGCEGPSHGGLPGASSEKGRKTYPTVKICNYEGPAKIEVDLVTHSDPP
p105 LEQPKQRGFRFRYVCEGPSHGGLPGASSEKNKKSYPQVKICNYVGPAKVIVQLVTNGKNI
      EQPKQRGFRFRY CEGPSHGGLPGASSEK  K YP VKICNY GPAK  V LVT

p49  RAHAHSLVGKQCSELGICAVSVGPKDMTAQFNNLGVLHVTKKNMMGTMIQKLQRQRLRSR
p100 RAHAHSLVGKQCSELGICAVSVGPKDMTAQFNNLGVLHVTKKNMMGTMIQKLQRQRLRSR
p105 HLHAHSLVGKHC-EDGICTVTAGPKDMVVGFANLGILHVTKKKVFETLEARM--------
       HAHSLVGK C E GIC V  GPKDM   F NLG LHVTKK    T

p49  PQGLTEA-----------------------EQRELEQEAKEL--------KKVMDLSIVR
p100 PQGLTEA-----------------------EQRELEQEAKEL--------KKVMDLSIVR
p105 ----TEACIRGYNPGLLVHPDLAYLQAEGGGDRQLGDREKELIRQAALQQTKEMDLSVVR
         TEA                        R L D  KEL          K MDLS VR

p49  LRFSAFLRASDGSFSLPLKPVTSQPIHDSKSPGASNLKISRMDKTAGSVRGGDEVYLLCD
p100 LRFSAFLRASDGSFSLPLKPVTSQPIHDSKSPGASNLKISRMDKTAGSVRGGDEVYLLCD
p105 LMFTAFLPDSTGSFTRRLEPVVSDAIYDSKAPNASNLKIVRMDRTAGCVTGGEEIYLLCD
      L F AFL  S GSF   L PV S  I DSK P ASNLKI RMD TAG V GG E YLLCD

p49  KVQKDDIEVRFYEDDENG--WQAFGDFSPTDVHKQYAIVFRTPPYHKMKIERPVTVFLQL
p100 KVQKDDIEVRFYEDDENG--WQAFGDFSPTDVHKQYAIVFRTPPYHKMKIERPVTVFLQL
p105 KVQKDDIQIRFYEEEENGGVWEGFGDFSPTDVHRQFAIVFKTPKYKDINITKPASVFVQL
      KVQKDDI  RFYE  ENG  W  FGDFSPTDVH Q AIVF TP Y    I  P  VF QL

p49  KRKRGGDVSDSKQFTYYPLVEDKEEVQRKRRKALPTFSQPFGGGSHMGGGSGGAAGGYGG
p100 KRKRGGDVSDSKQFTYYPLVEDKEEVQRKRRKALPTFSQPFGGGSHMGGGSGGAAGGYGG
p105 RRKSDLETSEPKPFLYYPEIKDKEEVQRKRQKLMPNFSDSFGGGSGAGAGGGGM---FGS
      RK     S  K F YYP   DKEEVQRKR K  P FS  FGGGS  G GGGG     G

p49  AGGG[EGVLMEGGVKVKEAVEEKNLGEAGRGLHACNPAFGRPRQADYLRSGVQDQLGQQR
p100 AGGGGSLGFFPSSLAYSPYQ-SGAGPMRCYPGGG-GGAQMAATVPSRDS--GEEAAEPSA
P105 GGGGGGTGSTGPGYSFPHYGFPTYGGITFHPGTTKSNAGMKHGTMDTESKKDPEGCDKSD
      GGGGG G          Y    G     PG     A M        S    E   S

p49  ETSSLLKIQTLAGHGGRRL*]
p100 PSRTP-------QCEPQAPEMLQRAREYNARLFGLA-----HAAPSPTRLLRHRG-----
p105 DKNTVNLFGKVIETTEQDQEPSEATVGNGEVTLTYATGTKEESAGVQDNLFLEKAMQLAK
       T           Q  E                A      A     L

p100 RRA--------------LLAGQRHLLTAQDENGDTPLHLAIIHGQTSVIEQIVYVIHHAQ
p105 RHANALFDYAVTGDVKMLLAVQRHLTAVQDENGDSVLHLAIIHLHSQLVRDLLEVTSGLI
     R A               LLA QRHL   QDENGD  LHLAIIH           V

p100 DLGVVNLTNHLHQTPLHLAVITGQTSVVSFLLRVGADPALLDRHGDSAMHLALRAGAGAP
p105 SDDIINMRNDLYQTPLHLAVITKQEDVVEDLLRAGADLSLLDRLGNSVLHLA--AKEGHD
          N  N L QTPLHLAVIT Q   VV   LLR GADL LLDR G S   HLA   A   G

p100 ELLRALLQSGAPAVPQLLHMPDFEGLYPVHLAVRARSPECLDLLVDSGAEVEATERQGGR
p105 KVLSILLKHKKAAL--LLDHPNGDGLNAIHLAMMSNSLPCLLLLVAAGADVNAQEQKSGR
       L  LL      A    LL  P    GL    HLA     S   CL LLV   GA V A E    GR

p100 TALHLATEMEELGLVTHLVTKLRANVNARTFAGNTPLHLAAGLGYPTLTRLLLKAGADIH
p105 TALHLAVEHDNISLAGCLLLEGDAHVDSTTYDGTTPLHIAAGRGSTRIAALLKAAGADPL
     TALHLA E      L    L      A V    T   G TPLH AAG G    L   LL   AGAD

p100 AENEEPLCPLPSPPTSDSDSDSEGPEKDTRSSFRGHTPLDLTCSTLVKTLLLNAAQNTME
p105 VENFEPLYDLDDS--WENAGEDEGVVPGT--------TPLDMATSWQVFDIL---NGKPYE
       EN EPL  L             EG     T        TPLD    S   V    L          E

p100 PPLTPPS-PAGPGLS-LGDTALQNLEQLLDGPEAQGSWAELAERLGLRSLVDTYRQTTSP
p105 PEFTSDDLLAQGDMKQLAEDVKLQLVKLLEIPDPDKNWATLAQKLGLGILNNAFRLSPAP
     P  T      A       L         L  LL   P      WA LA   LGL   L     R     P

p100 SGSLLRSYELAGGDLAGLLEALSDMGLEEGVRLLRGPETRDKLPSTEVKEDSAYGSQSVE
p105 SKTLMDNYEVSGGTVRELVEALRQMGYTEAIEVIQAASSPVKTTSQA--HSLPLSPASTR
     S  L    YE   GG     L EAL   MG   E                K   S                S

p100 QEAEKL-GPPPEPPGGLSHGHPQPQVTDLLPAPSPLPGPPVQRPHLFQILFNTPHPPLSW
p105 QQIDELRDSDSVCDTGVETSFRKLSFTESLTSGASLLTLN-KMPHDY-----GQEGPLE-
     Q    L          G              T  L      L         PH              PL

p100 DK*
p105 GKI*
       K
```

c

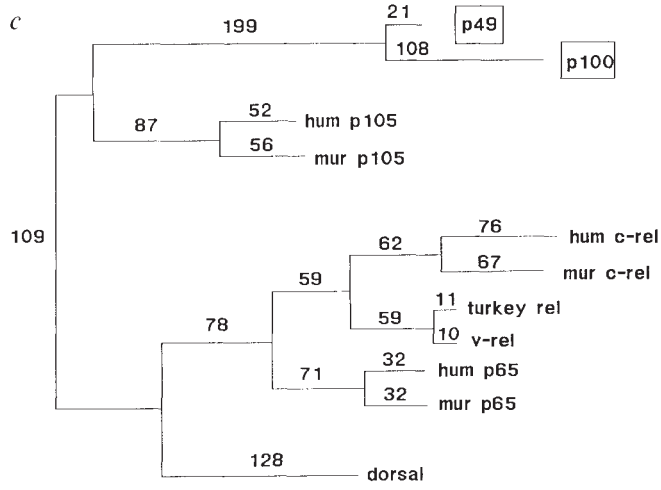


FIG. 2 Deduced amino-acid sequence of the p49 and p100 NF-κB cDNA clones, and similarity to other NF-κB/rel/dorsal genes. *a*, Schematic comparison of the major structural features of p49 and p100. *b*, The deduced amino-acid sequence of p49 and p100 and comparison to p105 NF-κB. The alternatively spliced C-terminal portion of p49 is indicated in brackets and is excluded from sequence comparison. The 1.9-kb p49 cDNA clone is presented with omission of a 220-bp intron present in one clone. The nucleotide sequence is submitted to Genbank/EMBL. *c*, Evolutionary relatedness of NF-κB, *rel* and *dorsal* genes determined using amino-acid sequence of the indicated genes with the PAUP program[15]. Branch lengths, the total number of amino-acid differences between nodes, are indicated. Comparison of p49 NF-κB sequence with human p105 NF-κB (ref. 3), the human c-*rel*[4,7,8], mouse c-*rel*[16], turkey c-*rel*[17], v-*rel*[17,18] and dorsal proteins[19,20] revealed 26% identity in the *rel* conserved region.

METHODS. A λgt11 library from the BJAB human B-lymphoid leukaemia[21] (from M. Kamps) was screened with a 700-bp fragment identical to the human c-*rel* gene using hybridization and washes at 50°C. Two independent isolates of ~1.9 and ~0.9 kb contained related sequences, with the 0.9-kb insert representing the 3′ end of the larger clone. The inserts were cloned into the bluescript SK+ plasmid (Stratagene) and sequenced using Sequenase (US Biochemicals). The p100 clone was isolated from a λgt11 cDNA library prepared from an Epstein-Barr virus-transformed B cell, GM 2471 (from D. Ginsburg).

(Fig. 1*b*), demonstrating that there are several species of p49 whose abundance varies among different cell types. For example, the larger 3.5 kb RNA was found in all the cells we examined and was predominant in YT T leukaemia cells, but both transcripts were present equally in the JY B-cell line (Fig. 1*b*, lanes 1 and 3).

To characterize the larger messenger RNA species, we isolated and sequenced additional cDNA clones. The clone for the shorter transcript contained a 1,341-base pair (bp) open reading frame, encoding a protein of predicted M_r ~49,100 (Fig. 2*a*, *b*); the longer clone encoded a protein of predicted M_r ~100,634. These sequences were identical through amino acid 374, into the glycine-rich putative hinge region, after which they diverged. The amino-acid sequence in the common N-terminal region was similar (26% identity) to that in p105 NF-κB, and in the *dorsal* and *rel* proteins (Fig. 2*a*), the greatest similarity being between p49/p100 and p105 (60% identity). One subregion contained conserved cysteine and histidine residues which do not conform to a classic zinc-finger motif. Interestingly, κB-binding activity depends on zinc[9], and, like the *tat-I* gene of HIV, may form an alternative structure which participates in dimerization or nucleic acid binding. The longer transcript contains repeated sequences in the C-terminal region which have homology with motifs in p105 and proteins encoded by cell cycle genes[3-6]. In this family, p100 NF-κB is most closely related to p105 NF-κB (41% identity), and p65 to c-*rel* protein (50% identity) (Fig. 2*c*).

To investigate the DNA-binding activity of p49, we used a prokaryotic expression system and an electrophoretic mobility shift assay. The predominant product of ~49K (Fig. 3*a*, lane 1) displayed specific κB binding activity (Fig. 3*b*, lane 2 versus

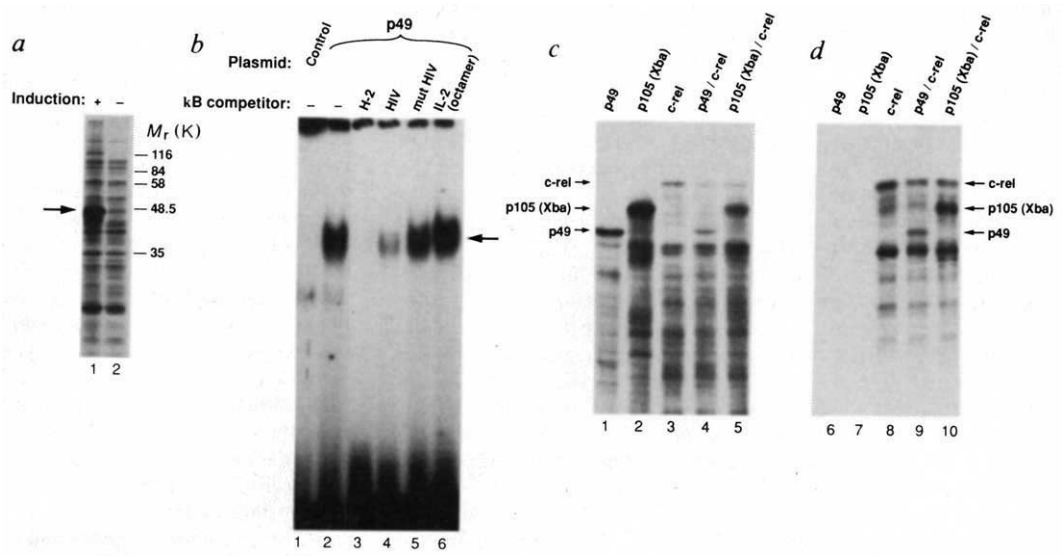


FIG. 3 κB binding and heterodimerization of the p49 cDNA clone. *a*, SDS-PAGE of recombinant p49 visualized by Coomassie brilliant blue staining in an induced (+) or uninduced *Escherichia coli* strain (−). The inducible protein band is indicated by an arrow. *b*, Electrophoretic mobility shift assays of bacterially expressed control (lane 1) or p49 protein (lanes 2–6). The assay was performed using a double-stranded ^{32}P-labelled oligonucleotide probe containing the H-2K^b enhancer κB site (0.1 ng). Unlabelled κB sites (100 ng) from the indicated enhancers were used for the competition reactions. Arrow denotes specific inducible complex. *c*, SDS-PAGE (lanes 1–5) and *d*, immunoprecipitation (lanes 6–10) of cDNAs translated *in vitro* using wheat-germ lysates. Immunoprecipitations were performed with a polyclonal antiserum raised against v-*rel* protein. [^{35}S]methionine-labelled protein extracts were resolved on a 10% SDS-polyacrylamide gel and visualized by autoradiography. Arrows indicate full-length p49, p105 (*Xba*I) and c-*rel* protein.

METHODS. p49 was expressed in *E. coli* using the pET system in the pLysE strain[22]. Protein extracts were prepared as described[10], except buffer A contained 0.05% Nonidet-P40, and glycerol was added to a concentration of 20% (v/v); electrophoretic mobility shift assay was performed by standard methods[23,24] using ~0.4 μg protein extract. The control extract was prepared from a p49 deletion mutant lacking the first 317 N-terminal amino acids.

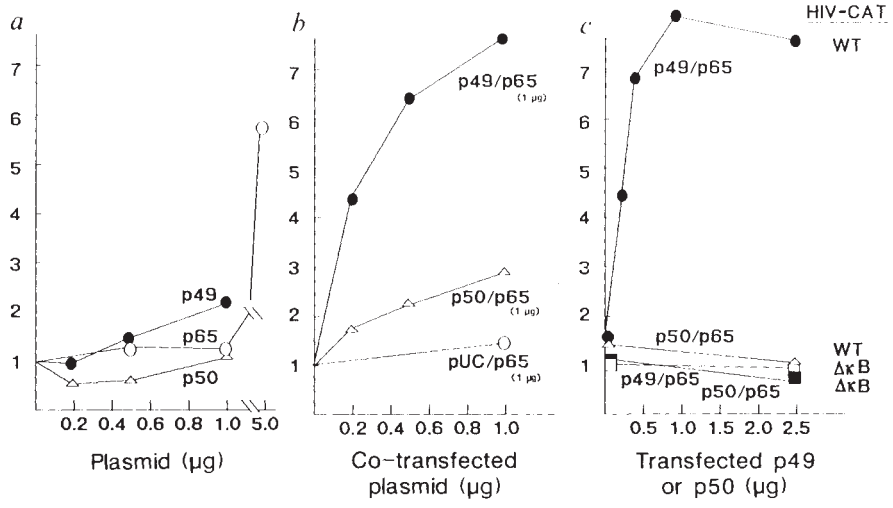


FIG. 4 Transcriptional activation and synergistic function of p49/p65 *in vivo*. Jurkat cells were transfected with a reporter gene containing 4 κB sites linked to CAT (ref. 23) and eukaryotic expression vectors encoding p49, p105 (*Rsa*I) or 'p50' and p65 alone (*a*) or in the indicated combinations with a suboptimal constant amount of p65 (*b*). Jurkat cells were also transfected with HIV-CAT or its κB mutant[24] with the indicated combinations and amounts of p49 or p50 (p105 (*Rsa*I)) expression plasmids (*c*) and a constant amount of p65 (1 or 2.5 μg). pUC 9 plasmid was added to make a total amount of 20 μg DNA per transfection. In separate experiments, no additional transactivation was observed with 5 μg of p49 or p105(*Rsa*I) alone in *a*, and similar results were observed with p105(*Rsa*I) and p105(*Xba*I) (data not shown). WT, Wild type.

METHODS. Full-length p105 cDNA was isolated from the GM 2471 human B-cell λgt11 library. The *Rsa*I truncation was generated by insertion of a stop codon at the position corresponding to amino acid 402, followed by an *Xho*I restriction site inserted by site-directed mutagenesis[24], which allowed deletion of the 3′ end of the p105 *Rsa* I-truncated gene. Mouse p65 cDNA was obtained by PCR with 70Z B-cell cDNA using primers from the sequence[25] corresponding to amino acids 1–546. All cDNAs had a *Hin*dIII restriction site and a consensus Kozak sequence (AAGCTTCACC*ATG*). The *Hin*dIII site allowed subcloning into a Rous sarcoma virus β globin expression vector[26] in which the β globin coding sequence had been removed by digestion with *Hin*dIII and *Bgl*II.

lane 1) and could compete with the H-2 and HIV κB sites, but not with a single-base-pair mutant of HIV or with an unrelated interleukin-2 octamer site (Fig. 3*b*, lanes 3 and 4 versus 5 and 6). We tested the interaction of p49 with other *rel* proteins by immunoprecipitation in a wheat-germ co-translation system: like p105 (*Xba*I), p49 also associated with c-*rel* protein (Fig. 3*d*).

To determine whether p49 interacts with other NF-κB/*rel* proteins to stimulate transcription, eukaryotic expression vectors were transfected into Jurkat cells. Transfection of p49 alone stimulates κB enhancer activity slightly; p65 at higher concentrations significantly increases κB-dependent transcription (Fig. 4*a*). Transfection of small amounts (≤1 μg) of either p49 or p65 causes minimal stimulation, but when co-transfected with each other, they act together to stimulate a κB reporter plasmid. This stimulation was more effective than the combination of 'p50' (p105(*Rsa*I)) and p65 (Fig. 4*b*), suggesting that p49 was more effective in cooperating with p65. When analysed with the HIV-chloramphenicol acetyl transferase (HIV-CAT) reporter, the p49/p65 combination, but not p50 (p105(*Rsa*I))/p65, stimulated HIV-CAT activity, an effect that required an intact κB regulatory element (Fig. 4*c*). Truncated p100 (48.5K form) gave a comparable stimulation, in contrast to full-length p100, which was inactive (data not shown), suggesting that differences in the *rel* conserved domain mediate this effect.

These findings demonstrate that κB-dependent transcription is regulated by the p49/100 gene products. The p65 and *rel* proteins are putative transcriptional activation subunits with intrinsic DNA-binding activity which associate with another DNA-binding subunit, previously thought to be a single gene product derived from p105. These findings suggest that p49/100 is an alternative DNA-binding subunit of NF-κB which acts with p65 to activate κB-dependent transcription. Many proteins can bind to κB-related sites, some of which are not NF-κB/*rel*-related (refs 10–12; and B. Adams, K. Leung, E. Hanley and G. J. N., manuscript submitted). Although some κB-binding proteins have been characterized[13,14], their identification is equivocal as they have related antigenic epitopes in their amino-terminal region. As with p49/100, the identification of these cDNAs will allow them to be analysed further. The interaction of p49/100 with other proteins and its other possible modes of regulation may provide additional mechanisms to regulate the transcription of HIV and different κB-containing genes. □

Received 8 May; accepted 8 July 1991.

1. Sen, R. & Baltimore, D. *Cell* **46,** 705–716 (1986).
2. Baeuerle, P. A. & Baltimore, D. *Genes Dev.* **3,** 1689–1698 (1989).
3. Kieran, M. *et al. Cell* **62,** 1007–1018 (1990).
4. Ghosh, S. *et al. Cell* **62,** 1019–1029 (1990).
5. Gilmore, T. *Cell* **62,** 841–843 (1990).
6. Bours, V., Villalobos, J., Burd, P. R., Kelly, K. & Siebenlist, U. *Nature* **348,** 76–80 (1990).
7. Brownell, E., O'Brien, S. J., Nash, W. G. & Rice, N. *Molec. cell. Biol.* **5,** 2826–2831 (1985).
8. Brownell, E., Mittereder, N. & Rice, N. R. *Oncogene* **4,** 935–942 (1989).
9. Zabel, U., Schreck, R. & Baeuerle, P. A. *J. biol. Chem.* **266,** 252 (1991).
10. Singh, H., LeBowitz, J. H., Baldwin, A. S. & Sharp, P. *Cell* **52,** 415–423 (1988).
11. Fan, C. M. & Maniatis, T. *Genes Dev.* **4,** 29–42 (1989).
12. Baldwin, A. S., LeClair, K. P., Singh, H. & Sharp, P. A. *Molec. cell. Biol.* **10,** 1406–1414 (1990).
13. Ballard, D. W. *et al. Cell* **63,** 803–814 (1990).
14. Franza, B. R., Josephs, S. F., Gilman, M. Z., Ryan, W. & Clarkson, B. *Nature* **330,** 349–355 (1987).
15. Quemada, H., Sieu, L. C., Siemieniak, D. R., Gonsalves, D., & Slightom, J. L. *J. Gen. Vir.* **71,** 1451–1460 (1990).
16. Grumont, R. J. & Gerondakis, S. *Oncogene Res.* **4,** 1–8 (1989).
17. Wilhelmsen, K. C., Eggleton, K. & Temin, H. M. *J. Virol.* **52,** 172–182 (1984).
18. Stephens, R. M., Rice, N. R., Hebsch, R. R., Bose, H. R. & Gilden, R. V. *Proc. natn. Acad. Sci. U.S.A.* **80,** 6229–6233 (1983).
19. Steward, R. *Cell* **59,** 1179–1188 (1989).
20. Steward, R. *Science* **238,** 692–694 (1987).
21. Kamps, M. P., Murre, C., Sun, X. H. & Baltimore, D. *Cell* **60,** 547–555 (1990).
22. Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorf, J. W. *Meth. Enzym.* **185,** 60–89 (1990).
23. Leung, K. & Nabel, G. *Nature* **333,** 776–778 (1988).
24. Nabel, G. & Baltimore, D. *Nature* **326,** 711–713 (1987).
25. Nolan, G. P., Ghosh, S., Liou, H. C., Tempst, P. & Baltimore, D. *Cell* **64,** 961–969 (1991).
26. Gorman, C., Padmanabhan, R. & Howard, B. H. *Science* **221,** 551–553 (1983).

ACKNOWLEDGEMENTS. This work resulted from an equal contribution by the first two authors. We thank D. Gschwend for typing assistance; E. Hanley for technical expertise; B. Adams and B.-y. Wu for northern blot filters. N. Rice for c-*rel* expression vectors and antibodies; A. Israel for analysing and sharing peptide sequence before publication; G. E. Griffin and D. Ginsburg for advice, discussion and help with computer graphics; and D. Siemieniak for help with computer analyses. This work was supported in part by a DFG Fellowship (R.M.S.), the British Medical Research Council's AIDS Directed Programme (C.S.D.), and a grant from the National Institutes of Health (G.J.N.). C.S.D. is also affiliated with St George's Hospital Medical School in London.

A protein-tyrosine phosphatase with sequence similarity to the SH2 domain of the protein-tyrosine kinases

Shi-Hsiang Shen, Lison Bastien, Barry I. Posner* & Pierre Chrétien

Section of Molecular Genetics, Biotechnology Research Institute, National Research Council of Canada, 6100 Royalmount Avenue, Montréal, Québec, Canada H4P 2R2
* Polypeptide Hormone Laboratory, Department of Medicine, Royal Victoria Hospital and McGill University, 3640 University Street, Montréal, Québec, Canada H3A 2B2

THE phosphorylation of proteins at tyrosine residues is critical in cellular signal transduction, neoplastic transformation and control of the mitotic cycle[1]. These mechanisms are regulated by the activities of both protein-tyrosine kinases (PTKs) and protein-tyrosine phosphatases (PTPases)[2]. As in the PTKs, there are two classes of PTPases: membrane associated, receptor-like enzymes[3–5] and soluble proteins[3,6,7]. Here we report the isolation of a complementary DNA clone encoding a new form of soluble PTPase, PTP1C. The enzyme possesses a large noncatalytic region at the N terminus which unexpectedly contains two adjacent copies of the Src homology region 2 (the SH2 domain) found in various nonreceptor PTKs[8] and other cytoplasmic signalling proteins[9–11]. As with other SH2 sequences, the SH2 domains of PTP1C formed high-affinity complexes with the activated epidermal growth factor receptor and other phosphotyrosine-containing proteins. These results suggest that the SH2 regions in PTP1C may interact with other cellular components to modulate its own phosphatase activity against interacting substrates. PTPase activity may thus directly link growth factor receptors and other signalling proteins through protein-tyrosine phosphorylation.

Screening of a human breast carcinoma (ZR-75-1) cDNA library with a LAR cDNA probe[4] under conditions of low stringency[5] yielded many positive phage clones. Southern blotting showed that four of these clones hybridized to each other under conditions of high stringency. Two clones with large cDNA inserts, PTP1C4 (2.3 kilobases (kb)) and PTP1C2 (2.1 kb), were chosen for further characterization. The sequence of the largest cDNA, PTP1C4, had an open reading frame potentially encoding a polypeptide of 609 amino acids (Fig. 1). The sequence of the second largest cDNA clone, PTP1C2, was identical to that of PTP1C4 but started at nucleotide 249 of PTP1C4. Northern blot analysis of RNA from ZR-75-1 cells[12] with the PTP1C cDNA as a probe detected a single transcript of about 2.4 kb. The size of this transcript agrees with the size of PTP1C4 cDNA insert.

To demonstrate that the cDNA PTP1C clone encodes an active PTPase, the PTP1C2 cDNA insert was subcloned into the *Bam*HI site of the expression vector pET-3c (ref. 13) to generate the plasmid pT7-PTP1C. *Escherichia coli* cells transformed with this plasmid were induced with isopropylthiogalactoside (IPTG) and cell lysates were assayed for PTPase activity. Although *E. coli* extracts carrying a control plasmid have no detectable PTPase activity, cells transformed with pT7-PTP1C have very high amounts of PTPase activity (Fig. 2). These results indicate that the deduced protein sequence in Fig. 1 encodes a functional PTPase.

Computer analysis of the amino-acid sequences of PTP1C did not find a sequence with the characteristics of a signal peptide in the N-terminal region. Neither was a characteristic transmembrane sequence found in the entire protein sequence. Accordingly, it is concluded that PTP1C is a cytosolic soluble PTPase. Analysis of the primary structure of PTP1C showed