The EMBO Journal vol.7 no.8 pp.2465-2473, 1988

Molecular cloning and characterization of the human *dbl* proto-oncogene: evidence that its overexpression is sufficient to transform NIH/3T3 cells

Dina Ron, Steven R.Tronick, Stuart A.Aaronson and Alessandra Eva

Laboratory of Cellular and Molecular Biology, National Cancer Institute, Building 37, Room 1E24, Bethesda, MD 20892, USA

Communicated by F.Blasi

We isolated cDNA clones representing the human *dbl* proto-oncogene transcript. Nucleotide sequence analysis revealed an open reading frame encoding a predicted protein of 925 amino acids. Using peptide antisera directed against specific proto-*dbl* peptides, a 115-kd protein was detected in COS cells transfected with an expression vector containing the entire coding region of proto-*dbl*. This mol. wt is consistent with that predicted from the open reading frame. We have previously shown that the *dbl* oncogene was generated by substitution of the 5′ portion of proto-*dbl* with an unrelated human sequence. In this study we show that this rearrangement resulted in the loss of the 497 amino-terminal codons of the *dbl* proto-oncogene. Under the influence of a strong promoter proto-*dbl* could readily transform NIH/3T3 cells but its transforming activity was less than that of the *dbl* oncogene driven by the same promoter. Proto-*dbl* overexpression is, therefore, sufficient to transform NIH/3T3 cells, but specific structural alterations of its coding region significantly enhance its transforming activity. No apparent similarity was detected between the predicted proto-*dbl* product and other known proto-oncogenes. However, a stretch of 300 amino acids within the N-terminal half of proto-*dbl* showed structural similarity to the intermediate filament vimentin. This region in proto-*dbl* contains a heptad repeat motif characteristic of an α-helical coiled-coil structure. Taken together, these findings indicate that the human proto-*dbl* represents a new class of cellular oncogenes that may be related to cytoskeletal elements of the cell.

Key words: proto-oncogene/*dbl*/overexpression/transformation/NIH/3T3 cells

Introduction

The *dbl* oncogene was isolated by transfection of NIH/3T3 cells with DNA of a human B-cell lymphoma (Eva and Aaronson, 1985). More recently, another example of the *dbl* oncogene was identifed in a human nodular, poorly differentiated lymphoma (designated NPDL-*dbl*) (Eva *et al.*, 1987). Comparison of the restriction maps of these independent *dbl* isolates with that of normal human DNA revealed that each arose as the result of recombinational events which replaced 5′ coding sequences of the *dbl* proto-oncogene with those of different human sequences (Eva *et al.*, 1987).

The *dbl* oncogene translational product is a 66-kd cytoplasmic phosphoprotein which is distributed equally between membrane and soluble fractions of *dbl*-transformed cells (Srivastava *et al.*, 1986). The predicted amino acid sequence of the *dbl* transcript shows no significant similarity to other known oncogene products or to any other sequence present in the NBRF data base (Eva *et al.*, 1988). Further computer-assisted analyses have offered no strong clues as to its function, but have indicated that pp66^{dbl} is markedly hydrophilic and possesses no characteristic membrane-spanning domain or pre-secretory signal peptide (Eva *et al.*, 1988). Thus, among known oncogenes, *dbl* represents a potentially new class. In an effort to decipher the normal function of the *dbl* proto-oncogene and study its mechanism of activation, we sought to isolate cDNA clones representing its coding sequence. Our findings provide insights into a possible novel function for this proto-oncogene as well as evidence that proto-*dbl* can be activated as a potent oncogene when expressed under the influence of a retroviral LTR.

Results

Cloning of human proto-dbl cDNA

In order to clone the proto-*dbl* cDNA, we analyzed poly(A)$^+$-containing RNAs from normal human tissues by utilizing RNA transfer analysis and hybridization with a probe (1b1-1) representing the 3′ 1.8-kb of the *dbl* oncogene cDNA (Eva *et al.*, 1987). A 5-kb transcript was detected in RNA from fetal brain and fetal adrenal gland. Based on this survey, aλgt11 human brain stem cDNA library was screened with the 1b1-1 probe. Four cDNA clones were isolated and designated pcl-29, pcl-38, pcl-33 and pcl-50. The relationship of these clones to each other, to 1b1-1 and to the biologically active *dbl* genomic clone (Eva and Aaronson, 1985) was determined and the results are shown in Figure 1. Pcl-29 and pcl-38 overlapped, and together contained 1.4 kb of proto-*dbl* cDNA. Pcl-33 and pcl-50 were identical with respect to their sizes and restriction maps, spanned 2.3 kb of the proto-*dbl* transcript and mapped just 5′ to pcl-29. No apparent overlap was detected between pcl-33 or pcl-50 and pcl-29 by cross-hybridization experiments. Digestion of either pcl-33 or pcl-50 with *Eco*RI gave rise to two fragments of 1.83 and 0.43 kb. Only the smaller fragment shared homology with *dbl* cDNA clone 1b1-1. The larger fragment, however, was *dbl*-related, since it was capable of hybridizing to the *dbl* genomic clone (*Eco*RI fragment no. 2, Figure 1).

Nucleotide sequence anlaysis of pcl-29, pcl-38 and pcl-33 revealed that they represented most of the open reading frame of proto-*dbl*. However, the entire coding region was not obtained due to the lack of overlapping sequences between pcl-33 and pcl-29. Alignment of their sequences with those of *dbl* clone 1b1-1 (Eva *et al.*, 1988) revealed a 7-nucleotide gap. To obtain a cDNA clone corresponding to this gap, we screened an oligo(dT)-primed human fetal adrenal cDNA

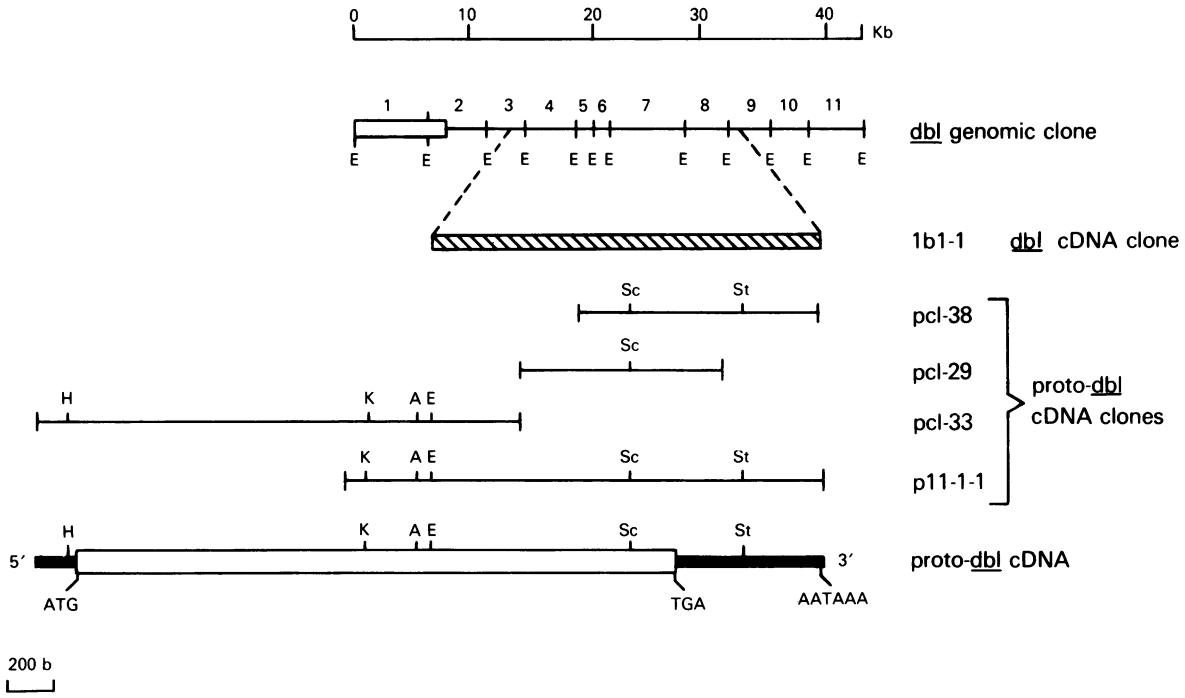

Fig. 1. Comparison of the structure of proto-*dbl* cDNA clones with a *dbl* genomic clone and *dbl* cDNA clone 1b1-1. The comparison was done by restriction endonuclease mapping of the various cDNA clones, by cross hybridization, as well as hybridization of each of the cDNA clones to *Eco*RI digested *dbl* genomic clone. Shown on top is a schematic representation of *dbl* genomic clone C14-1-2 (Eva and Aaronson, 1985; Eva *et al.*, 1987). The open box indicates the extent of the 5′ *dbl* oncogene rearrangement. The *dbl* cDNA clone 1b1-1 is indicated by the hatched box. The regions of the *dbl* genomic clone corresponding to clone 1b1-1 was determined by Southern blot analysis. Proto-*dbl* cDNA clones are shown below the *dbl* cDNA clone 1b1-1, and the schematic diagram of the composite proto-*dbl* cDNA clone is shown at the bottom. The solid lines indicate the 5′ and 3′ non-coding sequences and the large open box represents the coding region. The upper scale refers to the *dbl* genomic clone and the lower scale refers to the different cDNA clones.

library. Several clones were isolated, and their sequence analysis revealed that each contained a 7-nucleotide stretch that filled the gap between pcl-29 and pcl-33. A representative clone, designated p11-1-1, is shown in Figure 1. Also shown in this figure is a schematic diagram of the composite human proto-*dbl* cDNA.

Nucleotide sequence of the human proto-dbl cDNA and the predicted amino acid sequence of its translational product

The nucleotide sequence of proto-*dbl* cDNA and the deduced amino acid sequence of its translational product are shown in Figure 2. A potential initiation codon was found at nucleotide position 175. Two in-frame stop codons were located, respectively, 14 and 74 nucleotides upstream of this ATG. The next methionine codon was at nucleotide position 205. However, only the ATG codon at position 175 conformed to Kozak's rule for translation–initiation (Kozak, 1984,1987). Thus, it seemed most likely that this ATG functions as the translation initiation signal for the human proto-*dbl* protein. A single large open reading frame extending from this ATG to nucleotide position 2949 encoded a primary translational product of 925 amino acids.

The composite proto-*dbl* cDNA sequence was 3663 nucleotides long, and it contained at least a portion of the 5′ untranslated (UTS) region and the entire 3′ non-coding region. The 5′ UTS was 173 nucleotides in length and was A:T rich. The 3′ non-coding region was 697 nucleotides in length. A consensus polyadenylation signal, AATAAA, was located 23 nucleotides upstream of a short poly(A) tract. In addition, the 3′ non-coding region contained sequences of the AUUU(A) type which have been described as characteristic of unstable, short-lived mRNAs (Shaw and Kamen, 1987).

The predicted sequence of the human proto-*dbl* protein contained several post-translational modification sites. Two potential N-linked glycosylation sites were present at asparagine 462 and 673 respectively. In addition, a potential site for palmitylation was found at Cys 343 and was followed by three hydrophobic amino acids, Leu-Leu-Ala (Willumsen *et al.*, 1984). A putative site for serine phosphorylation by a cyclic AMP-dependent protein kinase was identified starting at amino acid 738, in agreement with the reported consensus sequence Arg-Arg-X-(X)-Ser-X (Creighton, 1984). As shown in Figure 3, the predicted proto-*dbl* translational product was markedly hydrophilic in its overall characteristics. Hydrophobic sequences indicative of a transmembrane domain or signal peptide (Perlman and Halvorson, 1983) were not apparent, suggesting that the *dbl* proto-oncogene protein is unlikely to be a secreted or an integral membrane protein.

Identification of the human proto-dbl translational product

We next sought to verify our sequencing data by expressing the proto-*dbl* cDNA and identifying its translational product with peptide antisera prepared against specific peptides from the predicted proto-*dbl* protein. To construct an expression vector, it was necessary to assemble the entire coding region from the individual cDNA clones. Thus, the

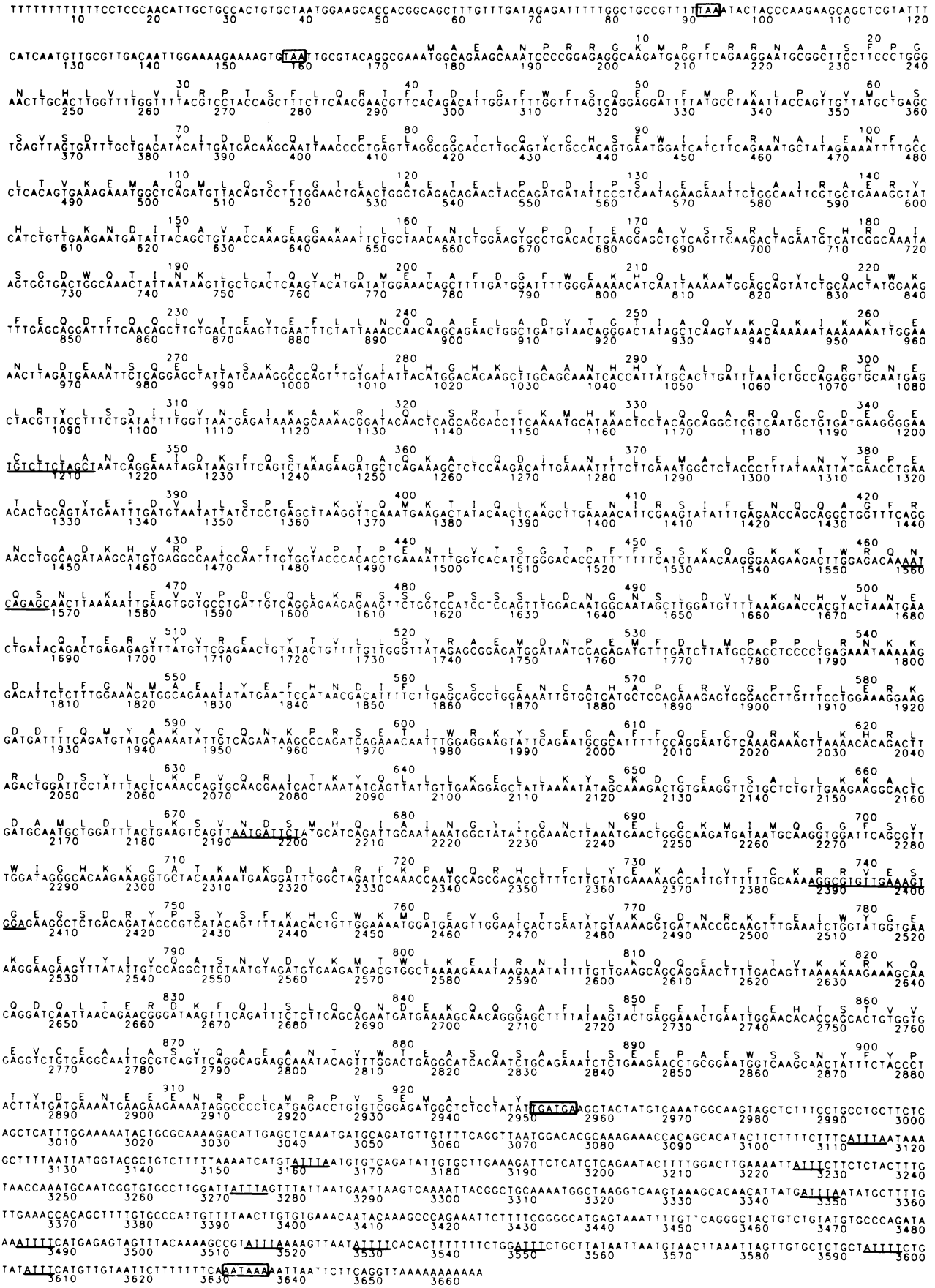

Fig. 2. Nucleotide sequence and deduced amino acid sequence of proto-*dbl* cDNA. The in-frame stop codons preceding the initiation codon, the two consecutive stop codons at the end of the reading frame and the polyadenylation signal are boxed. Potential N-linked glycosylation sites, putative palmitylation site and the consensus sequences for serine phosphorylation are underlined. The ATTT(A) sequences in the 3′ untranslated region are underlined as well. This figure was generated using the DRAW program (Shapiro and Senapathy, 1986).

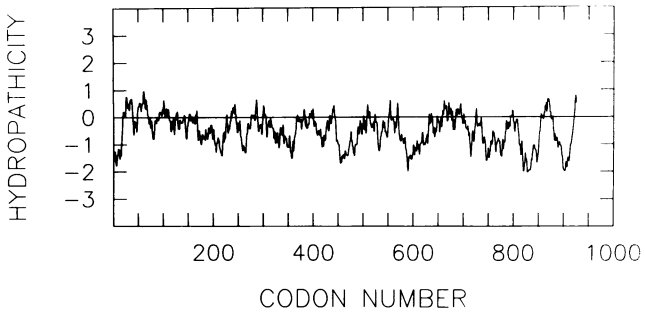

Fig. 3. Hydropathic profile of proto-*dbl* (Kyte and Doolittle, 1982). A window of 21 was used. Hydrophobic values are >0 and hydrophilic values are <0.

appropriate portions of pcl-38, pcl-33 and p11-1-1 were ligated in two intermediate plasmids (see Materials and methods). A fragment of the assembled cDNA clone spanning the *Hin*cII−*Sty*I sites was subcloned into the *Bam*HI site of the pcDV expression vector containing the SV40 early promoter region (Okayama and Berg, 1983). The cDNA insert in this construct, designated OB-S11-Cl1, contained 38 nucleotides of the 5′ UTS, the entire proto-*dbl* coding sequence and 238 nucleotides of the 3′ UTS.

OB-S11-Cl1 was transfected into COS-1 cells (Gluzman, 1981) by the calcium phosphate technique (Graham and van der Eb, 1973; Wigler *et al.*, 1979). Forty-eight hours post-transfection, the cells were metabolically labeled and subjected to immunoprecipitation analysis with antisera raised to peptides derived from codons 3−16 (designated anti-*dbl*-1) and 592−606 (designated anti-*dbl*-2) of the predicted proto-*dbl* sequence. The results are shown in Figure 4. Each antiserum detected a unique protein of 115 kd in cells transfected with the proto-*dbl* construct (panels A and B, lane 3) but not in mock-transfected cells (panels A and B, lane 1). This band was efficiently competed by preincubation of the antiserum with the corresponding peptide (lane 4 in both panels).

If the ATG codon at position 175 were to initiate the translation of human proto-*dbl*, the predicted mol. wt of this protein without any post-translational modifications would be 108 kd. The size of the proto-*dbl* protein identified in COS cells was 115 kd. These results suggest that in COS cells the proto-*dbl* protein does not undergo major post-translational modifications. In addition, our findings that anti-*dbl*-1 was able to recognize the gene product is consistent with the conclusion that the ATG at position 175 can function as the proto-*dbl* initiation codon.

Structural similarity of the predicted proto-dbl product to the rod domain of the intermediate filament vimentin

We searched the National Biomedical Research Foundation (NBRF) database (version 13) by using the FASTP program of Lipman and Pearson (1985) in order to determine the presence of any similarity between the proto-*dbl* amino acid sequence and those of known oncogenes. No significant matches were obtained. After searching the entire data base, we found that the sequence with the highest computer-assigned score (117) (compared to the mean value of 44.85) was that of hamster vimentin (Quax *et al.*, 1983), a member of the intermediate filament gene family. A statistical analysis of the alignment indicated that the match was 'probably

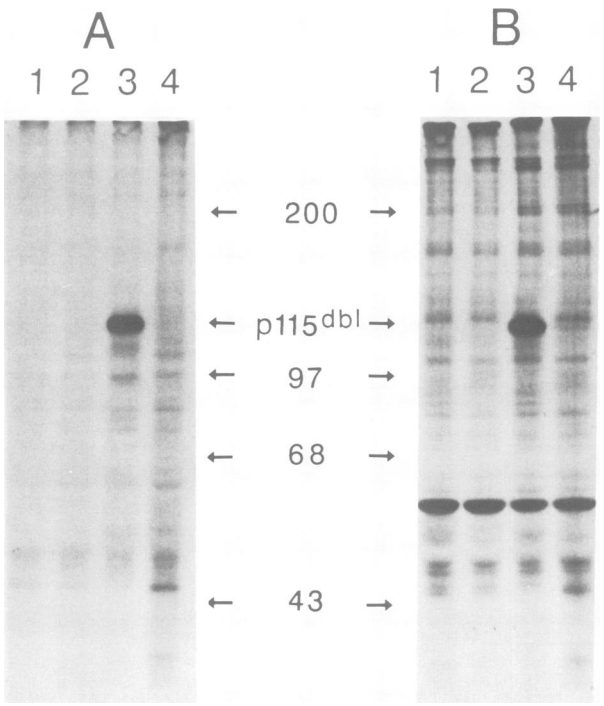

Fig. 4. Immunoprecipitation of the proto-*dbl* translational product from COS-1 cells. Mock-transfected (**lanes 1** and **2**) and OB-S11-Cl1-transfected (**lanes 3** and **4**) COS-1 cells were subjected to radio-immunoprecipitation with anti-*dbl*-1 (**panel A**) and anti-*dbl*-2 (**panel B**) respectively. **Lanes 1** and **3**, cell lysates were incubated with the indicated antisera alone; **lanes 2** and **4**, the antisera were preincubated with excess of the corresponding peptide prior to immunoprecipitation.

significant' (z value = 8.1) (Lipman and Pearson, 1985). The alignment in which 17% of the amino acids were identical extended for >300 residues and involved the rod domain of vimentin. When the comparison of the aligned region between proto-*dbl* and vimentin was performed on the basis of both chemically related and evolutionarily conserved amino acids (Lipman and Pearson, 1985), the similarity increased to 64%. We also observed that *Caenorhabditis elegans* myosin and rat cardiac muscle myosin heavy chain 1 were assigned scores of 103 and 97 respectively. The similarity with these proteins was in their rod domain as well.

The rod domains of intermediate filaments and myosin consist of α-helical coiled-coil structures. A prerequisite for the formation of a coiled-coil α-helix is the presence of heptad repeats, in which positions a and d usually contain apolar amino acid residues. Such repeats were present in the *dbl* sequence (Figure 5) and matched reasonably well with those of vimentin (Weber and Geisler, 1984). Secondary structure predictions of the proto-*dbl* product indicated that this region has the potential to form an α-helix (Chou and Fasman, 1978). Some heptads in proto-*dbl* were not completely regular. For example, several charged amino acids and two glycines were found in positions a and d (Figure 5). In addition, four proline residues in different positions were found in the region matching coil II of vimentin. However, the presence of such residues is not unusual and is thought to cause marginal stability in the coiled-coil by introducing breaks in the heptad (Cohen and Parry, 1986).

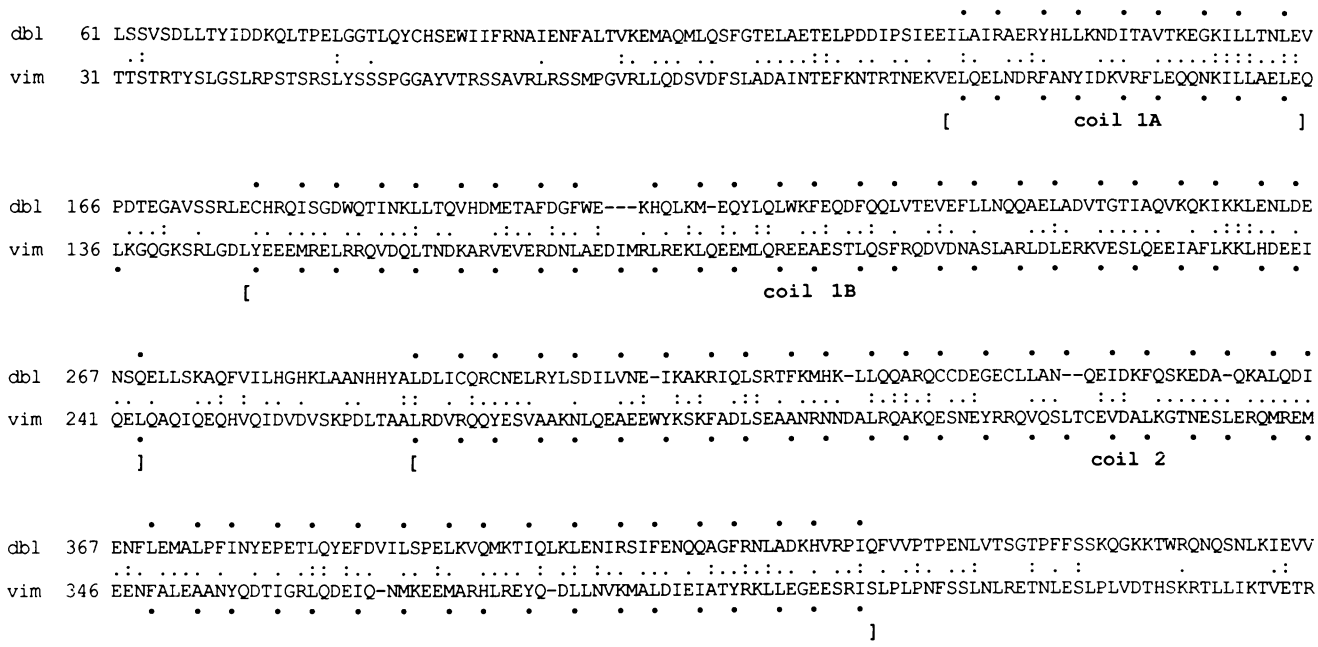

Fig. 5. Comparison of the amino acid sequences of human proto-*dbl* and the rod domain of the intermediate filament vimentin. The colons indicate amino acid identity and the single dots indicate evolutionarily conserved and chemically related amino acids. The large dots above and below the sequences indicate positions a and d in the heptad repeat of vimentin. The locations of the heptad repeats in hamster vimentin were taken from Weber and Geisler (1984).

Expression of proto-dbl mRNA in normal human tissues

In an effort to gain insight into proto-*dbl* function, we surveyed a wide variety of normal tissues for evidence of *dbl* expression by using Northern blotting analysis and hybridization with proto-*dbl* cDNA. We analyzed RNA from whole embryo, placenta and fetal brain, adrenal gland, liver and adenoid tissues. We also examined RNAs from adult human lung, kidney, skin, liver, testes, ovary, thymus, spleen, large intestine, tonsils, stomach, cardiac muscle and peripheral blood lymphocytes or mononuclear cells. Proto-*dbl* RNA species were detected only in fetal brain, fetal adrenal glands, as well as adult testes and ovaries (Figure 6). In addition to a common 5-kb proto-*dbl* transcript detected in each, adult testes and ovaries contained two additional bands of 3.3 and 1.9 kb. The lack of proto-*dbl* transcripts in other tissues was not due to low levels of RNA loaded on the gel, since the level of β-actin mRNA was similar in all samples analyzed (data not shown). These results suggest that the proto-*dbl* transcript is expressed and possibly processed in a highly tissue-specific manner.

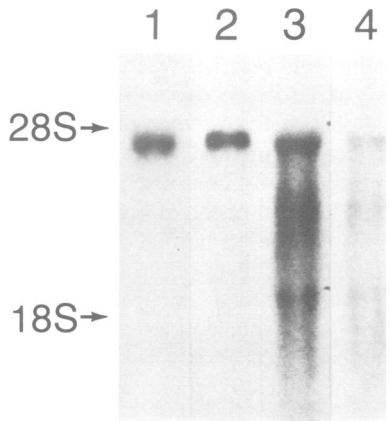

Fig. 6. Expression of the proto-*dbl* transcript in normal human tissues. A ^{32}P-labeled probe prepared from proto-*dbl* cDNA was hybridized to RNA extracted from normal human tissues. **Lane 1**, human fetal brain; **lane 2**, human fetal adrenal; **lane 3**, adult testes; **lane 4**, adult ovary.

Transforming potential of proto-dbl

It has been previously shown that some proto-oncogenes, such as c-H-*ras* (Chang *et al.*, 1982), c-*sis* (Gazit *et al.*, 1984) and c-*erb*B-2 (DiFiore *et al.*, 1987), can transform murine fibroblasts when their expression is driven by a strong promoter. In order to investigate whether human proto-*dbl* possesses this property, we introduced its cDNA into an expression vector in which transcription is controlled by Moloney murine leukemia virus long terminal repeats (Mo-MuLV LTR) (Cepko *et al.*, 1984). Previous studies indicate that this promoter is utilized very efficiently in NIH/3T3 cells (DiFiore *et al.*, 1987). The construct, designated LTR/proto-*dbl*, was transfected into NIH/3T3 cells, and its focus-forming activity was monitored. As shown in Table I, the proto-*dbl* construct induced foci of transformed cells at an efficiency of $4-7 \times 10^3$ focus-forming units (f.f.u.) per pmol of DNA, which was comparable to that of a viral LTR-driven human c-*sis* gene.

To analyze the growth properties of the proto-*dbl*-transformed NIH/3T3 cells, we determined their ability to display anchorage-independent growth and tumorigenicity in nude mice, properties which are known to correlate well with the malignant phenotype. Following LTR proto-*dbl* transfection and marker selection, a mass population of cells was suspended in semisolid medium or inoculated subcutaneously into nude mice. As shown in Table II, the colony-forming

efficiency of LTR/proto-*dbl* construct was comparable to that of NIH/3T3 cells transformed by LTR-driven human c-*sis*. LTR/proto-*dbl*-transfected cells were also tumorigenic in nude mice. All mice developed tumors within 2 weeks following inoculation of 10^5 cells, while mice inoculated with as many as 10^6 cells containing the expression vector alone were negative for tumors after 2 months. These results established that the human proto-*dbl* can be activated as a potent oncogene when its expression is driven by a strong promoter.

Comparison of the structure and transforming potential of proto-dbl and dbl oncogene cDNAs

We have previously shown that the *dbl* oncogene was generated by substitution of the 5′ portion of the *dbl* proto-oncogene with an unrelated human sequence (Eva *et al.*, 1987). In order to locate precisely the recombination site in *dbl* and to determine the extent of rearrangement, we compared the cDNA sequences of proto-*dbl* with that of the *dbl* oncogene (Eva *et al.*, 1988). A schematic diagram of both cDNAs is shown in Figure 7. The *dbl* oncogene cDNA represents a truncated version of its normal counterpart. The breakpoint was located at nucleotide position 1663, corresponding to amino acid 497 of the predicted proto-*dbl* protein. The nucleotide sequence from position 1664 through the 3′ end of proto-*dbl* was identical to that of the *dbl* oncogene, except for three nucleotide differences corresponding to nucleotide positions 2136, 2280 and 2831 of the *dbl* proto-oncogene sequence. The first two did not change the amino acid residue, whereas the third involved a C (proto-*dbl*) to T (*dbl*-oncogene) transition, resulting in a conservative change from *ala* in proto-*dbl* to *val* in the *dbl* oncogene (see Figure 7).

In contrast to the almost complete identity between their 3′ regions, the 5′ region of the *dbl* oncogene cDNA strikingly differed from that of its normal counterpart. The entire nucleotide sequence upstream of position 1663 in proto-*dbl* was deleted from the *dbl* oncogene cDNA and was substituted with 476 nucleotides of unrelated human sequences. These findings established that the *dbl* oncogene protein lacked the first 497 amino acids of the proto-*dbl* product and contained a new amino terminus of 50 amino acids derived from a different human locus.

To assess the contribution of *dbl* rearrangement to its activation as a transforming gene, we cloned its cDNA (Eva *et al.*, 1988) (Figure 7) into the same vector used to express proto-*dbl* cDNA in NIH/3T3 cells. Comparison of the focus-forming activity of both constructs upon transformation of NIH/3T3 cells revealed significant differences in their

Table I. Focus-forming activity of proto-*dbl* cDNA

DNA transfected	f.f.u./pmol DNA	neo^R colonies/ pmol DNA
LTR/proto-*dbl*-1	7×10^3	7×10^4
LTR/proto-*dbl*-2	4×10^3	4×10^4
LTR/c-*sis*	4×10^3	NT
pZIP-Neo SV(X)	$<1 \times 10^0$	1×10^5

Transfection assays were carried out by titration of each cloned DNA on recipient NIH/3T3 cultures. Foci were scored from 5 to 14 days post-transfection. Selection for neo^R expression was done by growing the cells in the presence of 375 mg/ml of G418. LTR/proto-*dbl*-1 and LTR/proto-*dbl*-2 are two clones of proto-*dbl* containing pZIP-Neo SV(X). NT: not tested.

Table II. Anchorage-independent growth and tumorigenicity of proto-*dbl* transformed cells

Donor DNA[a]	Colony-forming efficiency in agar[b] (%)	Tumorigenicity (cells inoculated[c])		
		10^6	10^5	10^4
LTR/proto-*dbl*	30	4/4	4/4	0/4
LTR/c-*sis*	30	4/4	3/4	0/4
pZIP-Neo SV(X)	<0.01	0/4	0/4	0/4

[a]1.5×10^5 NIH/3T3 cells were transfected with 1 μg of cloned DNA as described in Materials and methods.
[b]Cell suspensions were plated at 10-fold serial dilutions in 0.33% soft agar medium containing 10% calf serum. Colonies comprising >100 cells were scored at 14 days.
[c]NFR nude mice were inoculated subcutaneously with 10^6, 10^5 or 10^4 transfected cells. Tumorigenicity is expressed as the ratio of mice with tumors versus the number of mice inoculated. Tumors developed within 17 days.

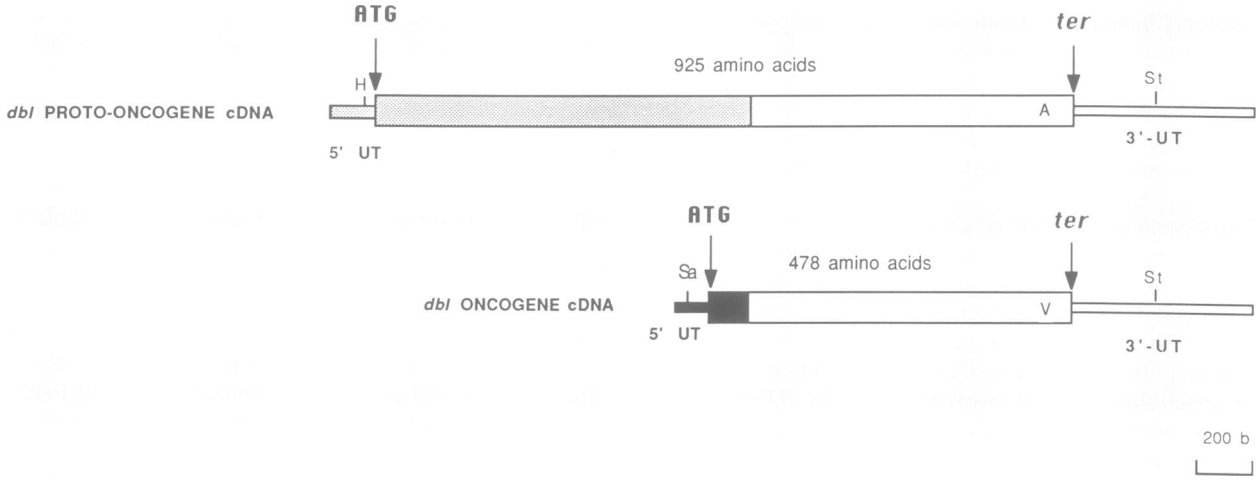

Fig. 7. Comparison of the structures of *dbl* proto-oncogene and *dbl* oncogene cDNAs. The open reading frames (ATG to *ter*) and 5′ and 3′ untranslated (UT) regions are indicated. The white boxes correspond to sequences shared by each cDNA. The single amino acid difference within this region (A−V) is shown. The stippled and black areas denote sequences unique to proto-*dbl* and *dbl* respectively. Restriction sites utilized to construct expression vectors are also shown (H = *Hin*cII, Sa = *Sac*II, St = *Sty*I).

transforming potencies. While proto-*dbl*-induced foci were first visualized between 10 and 14 days post-transfection, those induced by the *dbl* oncogene were detected as early as 6 days post-transfection. Moreover, the focus-forming activity of the oncogene was $1-5 \times 10^5$ f.f.u./pmol DNA, an efficiency 20−50 times higher than that of its normal counterpart (Table I). The level of expression of the 115-kd proto-*dbl* and 66-kd *dbl* proteins were similar in their respective transformants as judged by immunoprecipitation and densitometry (data not shown).

When transformants induced by each DNA construct were examined for anchorage-independent growth, their colony forming efficiencies were similar. However, colonies containing the *dbl* oncogene grew much more rapidly and achieved larger size. Finally, 10^3 *dbl*-transformed cells were sufficient to produce tumors in all injected nude mice within 2 weeks of inoculation, whereas 10^5 *dbl* proto-oncogene-transformed cells were required for 100% tumor incidence (Table II). Taken together, our results demonstrate that while overexpression is sufficent to activate proto-*dbl* as a transforming gene, gross structural alterations affecting its 5′ coding sequence can significantly enhance its transforming potential.

Discussion

In the present study, we report the molecular cloning and nucleotide sequence of the human *dbl* proto-oncogene cDNA. This gene is the progenitor of the *dbl* oncogene which was first revealed by DNA transfection analysis (Eva and Aaronson, 1985). The subset of cellular genes that can be activated as transforming genes include those that participate in growth factor-mediated cell proliferative pathways, such as growth factors themselves (Doolittle *et al.*, 1983; Waterfield *et al.*, 1983; Delli-Bovi *et al.*, 1987) and growth factor receptors (Downward *et al.*, 1984; Schechter *et al.*, 1984; Sherr *et al.*, 1985; King *et al.*, 1985). The products of *ras* genes are associated with the cytoplasmic membrane and, because of their resemblance to G proteins, are thought to play critical but as yet undetermined roles in signal transduction (Berridge and Irvine, 1984). Proto-oncogenes which encode protein kinases localized to the cytoplasm are also thought to participate in this process (Hunter and Cooper, 1985). Finally, certain proto-oncogene products which are localized to the nucleus, appear to be rapidly induced in response to growth factors (Kelly *et al.*, 1983; Greenberg and Ziff, 1984; Muller *et al.*, 1984) and may regulate expression of other genes by direct interactions with DNA (Kingston *et al.*, 1985). Our computer analyses of the predicted proto-*dbl* amino acid sequence revealed that it is not similar to any known oncogene. Thus, if proto-*dbl* is a component of the pathway in which growth factors induce cell proliferation, it may function in a novel way.

The proto-*dbl* protein was predicted to be hydrophilic and it lacked stretches of hydrophobic amino acids characteristic of secreted or integral membrane proteins. A potentially important feature was a stretch of ~300 amino acids in its N-terminal half, which showed some similarity to filamentous proteins. The highest score was with the rod domain of vimentin, a member of the intermediate filaments (IFs) family. This domain in all known IFs consists of ~310 amino acids, showing a seven-residue repeat pattern, which is required to form an α-helical coiled-coil structure. This rod domain is typically subdivided by non-helical spacers into three coils, termed coils 1a, 1b and II. It is highly conserved among IFs whereas the heads and tails of these proteins vary considerably (Steinert *et al.*, 1985; Franke, 1987). The identity within the rod domain of each type of IF can approach 70%, but the value can drop to 30% when different types are compared (Osborn and Weber, 1986). The most well-conserved consensus sequences are present in all IFs early in coil 1a and in the end of coil II (Steinert *et al.*, 1985).

The amino acid identity between the predicted proto-*dbl* protein and vimentin was 17% but the match did not correspond to the consensus type sequences mentioned above. However, when the match was done on the basis of chemically related and evolutionarily conserved amino acids, the similarity increased to 64%. When other IF sequences, such as GFAP (glial fibrillary acidic protein) (Weber and Geisler, 1981) or the nuclear lamins A and C (McKeon *et al.*, 1986) were compared in the same manner, the degree of similarity was also increased (between 35 and 50% within the different coils). Although these findings do not allow us to conclude that proto-*dbl* is a member of the IF gene family, they do suggest that most of the N-terminal half of the *dbl* proto-oncogene product could be arranged in a manner similar to that of the rod domain of IFs. The presence of heptad repeat motif and its predicted secondary structure within this region of similarity with IFs support this idea. The availability of the proto-oncogene cDNA clone will now make it possible to determine whether or not proto-*dbl* plays a structural role or has any relationship with cytoskeletal elements of the cells.

We observed that the *dbl* proto-oncogene exhibited a high degree of tissue-specific expression. Its transcript was not detected in a wide array of normal tissues and was found only in gonads, adrenal tissue and brain. The high degree of tissue specificity observed for the proto-*dbl* transcript argues against proto-*dbl* playing a generalized housekeeping function.

A number of proto-oncogenes have been shown to acquire transforming activity when overexpressed (Blair *et al.*, 1981; De Feo *et al.*, 1981; Chang *et al.*, 1982; Gazit *et al.*, 1984; DiFiore *et al.*, 1987). By expressing the normal *dbl* coding sequence under the influence of a strong promoter, we established that the normal gene itself possesses transforming potential. In fact, normal *dbl* under the control of a viral LTR influence was as efficient as some other known proto-oncogenes in inducing transformation of NIH/3T3 cells (Gazit *et al.*, 1984; DiFiore *et al.*, 1987).

Several oncogenes have been shown to acquire increased transforming efficiency as a result of truncations that affect the 5′ regions of their coding sequences. For example, removal of the putative ligand-binding domain of the *erb*B-2 gene increases its transforming potency (DiFiore *et al.*, 1987). The *raf* proto-oncogene bears structural similarities to protein kinase C, whose N-terminal domain is known to exert negative regulatory influence on the activity of its serine/threonine kinase domain (Inoue *et al.*, 1977; Ishikawa *et al.*, 1986). Truncations of the 5′ region of c-*raf* which occur naturally in tumors or as a result of the transfection process have been shown to activate this gene (Fukui *et al.*, 1987; Ishikawa *et al.*, 1987; Shimizu *et al.*, 1987).

The independent activation of *dbl* oncogenes in two cases involved genetic rearrangements that truncated the *dbl* gene

at its 5′ end and replaced 5′ non-coding and coding information with independently acquired sequences. The products of these independently activated *dbl* oncogenes are 66 and 76 kd respectively (Eva *et al.*, 1987), compared to the 115 kd established in our present studies for the normal *dbl* product expressed in COS cells. The deletion of proto-*dbl* that contributed to activation of the prototype *dbl* oncogene included the entire region encoding its predicted coiled-coil structure. From restriction mapping and protein analysis, the NDPL *dbl* oncogene appears to have lost most of this region as well (Eva *et al.*, 1987). The prototype *dbl* oncogene was shown in our present studies to be significantly more active as a transforming gene than its normal counterpart when both were expressed under the influence of the same promoter. Since in both cases of *dbl* activation, the 5′ acquired sequences differ from each other (Eva *et al.*, 1987), it is likely that the deletion of proto-*dbl* 5′ coding sequences plays a critical role in increasing the transforming potency of the gene. If so, the 5′ end of proto-*dbl* may exert a negative regulatory influence on its normal biological function as well.

Our previous findings have indicated that the rearrangements leading to independent activation of *dbl* oncogenes may have occurred during the DNA transfection process rather than within the tumors themselves (Eva *et al.*, 1987). Thus, evidence for a direct contribution of *dbl* to the development of naturally occurring tumors remains to be established. Our present findings that overexpression of the normal *dbl* gene can be sufficient to activate its transforming potential makes it worthwhile to search for evidence of proto-*dbl* overexpression in tumor cells. Such evidence, as well as findings of genetic rearrangements of the type described for the *dbl* oncogene, would support the role for *dbl* in naturally occurring malignancies.

Materials and methods

Materials

^{32}P-Labeled nucleotides were purchased from New England Nuclear and Amersham. Restriction enzymes and enzymes used in cDNA cloning and plasmid constructions were obtained from New England Biolabs, Boehringer-Mannheim and Pharmacia. Oligonucleotides were synthesized by the phosphoramidite method (Caruthers *et al.*, 1982).

General methods

Preparation of plasmid DNA, RNA extraction and fractionation, screening of λ phage and plasmid cDNA libraries, DNA blotting and hybridization were carried out using standard procedures as described by Maniatis *et al.* (1982). RNA blot hybridization was carried out as previously described (Thomas, 1980). DNA probes were prepared by nick-translation by labeling with polynucleotide kinase (Maniatis *et al.*, 1982), or by random hexanucleotide priming (Feinberg and Vogelstein, 1983). DNA sequencing was done by the dideoxy chain termination method of Sanger *et al.* (1977).

cDNA library screening

A bacteriophage λgt11 human brain stem cDNA library (generously provided by Dr Robert Lazzarini) was screened by plaque hybridization with ^{32}P-labeled *dbl* cDNA clone 1b1-1 (Eva *et al.*, 1988) in 50% formamide, 5 × SSC, 2.5 × Denhardt's solution, 10% dextran sulfate, 25 mM sodium phosphate buffer, pH 6.5, and 150 mg/ml of salmon sperm DNA at 42°C. Filters were washed twice at 55°C with 0.1 × SSC 0.1% SDS for 20 min each. An oligo(dT)-primed human fetal adrenal cDNA library (Yamamoto *et al.*, 1984, kindly provided by Dr P.C.White with the permission of Dr D.W.Russell) was screened with a probe corresponding to nucleotides 1835–2255 of proto-*dbl* derived from clone pcl-33 (Figure 1).

Construction of expression vectors

Expression vectors containing the entire coding region of proto-*dbl* were constructed as follows: first, the inserts of pcl-33, p11-1-1 and pcl-38 were ligated via their overlapping *Acc*I and *Sca*I sites (Figure 1) in pUC18 plasmid. This plasmid, designated p18-nd-3.6, contained the entire proto-*dbl* cDNA. The *Hinc*II–*Sty*I fragment was isolated from this plasmid, and the ends were filled in with Klenow fragment, and then *Bam*HI linkers were added for insertion into the *Bam*HI sites of the expression vectors pZIP-Neo SV(X) (Cepko *et al.*, 1984) and pcDV (Okayama and Berg, 1983). A fragment (*Sac*II–*Sty*I) containing the *dbl* oncogene-coding sequence (Eva *et al.*, 1988) was similarly manipulated and inserted into the pZIP-Neo SV(X) expression vector.

Immunoprecipitation

Peptide antisera were obtained by injecting rabbits with synthetic peptides corresponding, respectively, to amino acids 3–16 and 592–606 of the predicted proto-*dbl* gene product (Sutcliffe *et al.*, 1980; Walter *et al.*, 1980). Cultures of NIH/3T3 transfectants were labeled with [^{35}S]methionine plus [^{35}S]cysteine upon reaching 90% confluence and then lysed with 1 ml of lysis buffer. Clarified extracts were immunoprecipitated using the indicated antisera, recovered on protein A/Sepharose beads and resolved by SDS/7% PAGE, as described by Srivastava *et al.* (1986).

DNA transfection assay

DNA transfer into NIH/3T3 or COS cells was done by the calcium phosphate precipitation method as described elsewhere (Graham and van der Eb, 1973; Wigler *et al.*, 1979). The transfected cells were grown in DMEM supplemented with 5% calf serum in the presence or absence of 375 mg/ml of G418.

Acknowledgements

We gratefully acknowledge Dr S.Srivastava for help in preparing the peptide antisera, Drs R.Lazzarini, P.C.White and D.W.Russell for providing cDNA libraries and G.Vecchio and M.Pech for useful discussion. We also wish to thank C.Pennington and R.Trimmer for excellent technical assistance.

References

Berridge,M.J. and Irvine,R.F. (1984) *Nature*, **312**, 315–321.
Blair,D.G., Oskarsson,M., Wood,T.G., McClements,W.L., Fischinger,P.J. and Vande Woude,G.F. (1981) *Science*, **212**, 941–943.
Caruthers,M.H., Beaucage,S.L., Becker,C., Efcavitch,W., Fisher,E.F., Galluppi,G., Goldman,R., de Haseth,P., Martin,F., Mattencci,M. and Stabinsky,Y. (1982) *Genet. Eng.*, **4**, 1–13.
Cepko,C.L., Roberts,B.E. and Mulligan,R.C. (1984) *Cell*, **37**, 1053–1062.
Chang,E.H., Furth,M.E., Scolnick,E.M. and Lowy,D.R. (1982) *Nature*, **297**, 479–483.
Chou,P.Y. and Fasman,G.D. (1978) *Adv. Enzymol.*, **47**, 45–148.
Cohen,C. and Parry,D.A. (1986) *Trends Biochem. Sci.*, **11**, 245–248.
Creighton,T.E. (1984) *Proteins: Structural and Molecular Principles*. W.H.Freeman, New York, pp. 83–84.
De Feo,D., Gonda,M.A., Young,H.A., Chang,E.H., Lowy,D.R., Scolnick, E.M. and Ellis,R.W. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 3328–3332.
Delli-Bovi,P., Curatola,A.M., Kern,F.G., Greco,A., Ittmann,M. and Basilico,C. (1987) *Cell*, **50**, 729–737.
DiFiore,P.P., Pierce,J.H., Kraus,M.H., Segatto,O., King,C.R. and Aaronson,S.A. (1987) *Science*, **237**, 178–181.
Doolittle,R.F., Hunkapiller,M.W., Hood,L.E., Devare,S.G., Robbins,K.C., Aaronson,S.A. and Antoniades,H.N. (1983) *Science*, **221**, 275–277.
Downward,J., Yarden,Y., Mayes,E., Scrace,G., Totty,N., Stockwell,P., Ullrich,A., Schlessinger,J. and Waterfield,M.D. (1984) *Nature*, **307**, 521–527.
Eva,A. and Aaronson,S.A. (1985) *Nature*, **316**, 273–275.
Eva,A., Vecchio,G., Diamond,M., Tronick,S.R., Ron,D., Cooper,G.M. and Aaronson,S.A. (1987) *Oncogene*, **4**, 355–360.
Eva,A., Vecchio,G., Rao,C.D., Tronick,S. and Aaronson,S.A. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 2061–2065.
Feinberg,A.P. and Vogelstein,B. (1983) *Anal. Biochem.*, **132**, 6–13.
Franke,W.W. (1987) *Cell*, **48**, 3–4.
Fukui,M., Yamamoto,T., Kawai,S., Mitsunobu,F. and Toyoshima,K. (1987) *Mol. Cell. Biol.*, **7**, 1776–1781.
Gazit,A., Igarashi,H., Chiu,I.-M., Srinivasan,A., Yaniv,A., Tronick,S.R., Robbins,K.C. and Aaronson,S.A. (1984) *Cell*, **39**, 85–97.
Gluzman,Y. (1981) *Cell*, **23**, 175–182.
Graham,F.L. and van der Eb,A.J. (1973) *Virology*, **52**, 456–467.
Greenberg,M.E. and Ziff,E.B. (1984) *Nature*, **311**, 433–438.
Hunter,T. and Cooper,J.A. (1985) *Annu. Rev. Biochem.*, **54**, 897–930.

Inoue,M., Kishimoto,A., Takay,Y. and Nishizuka,Y. (1977) *J. Biol. Chem.*, **252**, 7610–7616.
Ishikawa,F., Takaku,F., Nagao,M. and Sugimura,T. (1986) *Jpn. J. Cancer Res.*, **77**, 1183–1187.
Ishikawa,F., Takaku,F., Nagao,M. and Sugimura,T. (1987) *Mol. Cell. Biol.*, **7**, 1226–1232.
Kelly,K., Cochran,B.H., Stiles,C.D. and Leder,P. (1983) *Cell*, **35**, 603–610.
King,C.R., Kraus,M.H. and Aaronson,S.A. (1985) *Science*, **229**, 974–976.
Kingston,R.E., Baldwin,A.S. and Sharp,P.A. (1985) *Cell*, **41**, 3–5.
Kozak,M. (1984) *Nucleic Acids Res.*, **12**, 857–872.
Kozak,M. (1987) *Nucleic Acids Res.*, **15**, 8125–8132.
Kyte,J. and Doolittle,R.F. (1982) *J. Mol. Biol.*, **157**, 105–132.
Lipman,D.J. and Pearson,W.R. (1985) *Science*, **227**, 1435–1441.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
McKeon,F.D., Kirschner,M.W. and Caput,D. (1986) *Nature*, **319**, 463–468.
Muller,R., Bravo,R., Burckhardt,J. and Curran,T. (1984) *Nature*, **312**, 716–720.
Okayama,H. and Berg,P. (1983) *Mol. Cell. Biol.*, **3**, 280–285.
Osborn,M. and Weber,K. (1986) *Trends Biochem. Sci.*, **11**, 469–472.
Perlman,D. and Halvorson,H.D. (1983) *J. Mol. Biol.*, **167**, 391–409.
Quax,W., Egberts,W.V., Hendriks,W., Quax-Jenken,Y. and Bloemendal,H. (1983) *Cell*, **35**, 215–223.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Schechter,A.L., Stern,D.F., Vaidyanathan,L., Decker,S.J., Drebin,J.A., Greene,M.I. and Weinberg,R.A. (1984) *Nature*, **312**, 513–516.
Shapiro,M.B. and Senapathy,P. (1986) *Nucleic Acids Res.*, **14**, 65–73.
Shaw,G. and Kamen,R. (1987) *Cell*, **46**, 659–667.
Sherr,C.J., Rettenmier,C.W., Sacca,R., Roussel,M.F., Look,A.T. and Stanley,E.R. (1985) *Cell*, **41**, 665–676.
Shimizu,K., Nakatsu,Y., Oh-uchida,M., Nomoto,S. and Sekiguchi,M. (1987) *Gan To Kagaku Ryoho*, **14**, 2140–2146.
Srivastava,S.K., Wheelock,R.H.P., Aaronson,S.A. and Eva,A. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 8868–8872.
Steinert,P.M., Steven,A.C. and Roop,D.R. (1985) *Cell*, **42**, 411–419.
Sutcliffe,J.G., Shinnick,T.M., Green,N., Liu,F.-T., Niman,H.L. and Lerner,R.A. (1980) *Nature*, **187**, 801–805.
Thomas,S.H. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5201–5205.
Walter,G., Scheidtmann,K.H., Carbone,A., Landano,A.P. and Doolittle,R.F. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5197–5200.
Waterfield,M.D., Scrace,G.T., Whittle,N., Stroobant,P., Johnsson,A., Wasteson,A., Westermark,B., Heldin,C.H., Huang,J.S. and Deuel,T.F. (1983) *Nature*, **304**, 35–39.
Weber,K. and Geisler,N. (1984) *Cancer Cells*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 153–159.
Wigler,M., Pellicer,A., Silverstein,S., Axel,R., Urlang,G. and Chasin,L. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 1373–1376.
Willumsen,B.M., Norris,K., Papageorge,A.G., Hubbert,N.L. and Lowy,D.R. (1984) *EMBO J.*, **3**, 2281–2285.
Yamamoto,T., Davis,C.G., Brown,M.S., Schneider,W.J., Casey,M.L., Goldstein,J.L. and Russell,D.W. (1984) *Cell*, **39**, 27–38.

Received on April 22, 1988