The EMBO Journal vol.3 no.10 pp.2355–2362, 1984

Sequencing of laminin B chain cDNAs reveals C-terminal regions of coiled-coil alpha-helix

Denise P. Barlow, N. Michael Green[1], Markku Kurkinen[2] and Brigid L.M. Hogan

Imperial Cancer Research Fund, Burtonhole Lane, Mill Hill, London NW7 1AD, and [1]National Institute for Medical Research, Mill Hill, London NW7 1AA, UK

[2]Present address: Department of Biochemistry, Rutgers Medical School, Piscataway, NJ 08854, USA

Communicated by D.A. Rees

cDNAs for laminin B chains have been isolated from a parietal endoderm cDNA library in pUC8 and pUC9. Identification is based on: (i) ability to direct the synthesis in *Escherichia coli* of polypeptides carrying laminin antigen determinants, (ii) *in vitro* translation of hybrid selected mRNA, and (iii) hybridization to high mol. wt. RNA differentially expressed in cells synthesizing large amounts of laminin. The plasmid pPE9 hybrid selects mRNA for the B2 (mol. wt. 185 000) chain and provides 217 residues of C-terminal amino acid sequence. The plasmids pPE386 and 49 both hybrid select mRNAs for the B1a (mol. wt. 205 000) and B1b (mol. wt. 200 000) chains. These two cDNAs are identical over much of their sequence, but pPE386 includes 133 nucleotides of 3′ non-coding sequence and a poly(A) tail. Together they provide 495 residues of C-terminal amino acid sequence. Analysis of the predicted sequences reveals a striking heptad repeat, with a high probability that residues *a* and *d* are hydrophobic. Such a repeat is typical of the coiled-coil alpha-helices found in proteins such as myosin, tropomyosin and desmin (2-stranded) and fibrinogen (3-stranded).

Key words: laminin/alpha-helix/cDNA/amino acid sequence

Introduction

Laminin is a large and complex glycoprotein present as a major component of all basement membranes (Timpl *et al.*, 1979; Foidart *et al.*, 1980; Martinez-Hernandez and Chung, 1984; Hogan *et al.*, 1984). It is synthesized by a variety of cell types, including epithelial cells, myoblasts, and astrocytes in primary culture (for review see Yamada, 1983; Kuhl *et al.*, 1982; Liesi *et al.*, 1983). Laminin is also the first extracellular matrix component synthesized by the mouse embryo, initially in the unfertilized egg and then from the four-cell morula stage onwards (Leivo *et al.*, 1980; Cooper and MacQueen, 1983; Wu *et al.*, 1983). *In vitro* studies have shown that laminin promotes the attachment and spreading of epithelial cells and hepatocytes and the outgrowth of neurites from neurones (Terranova *et al.*, 1980; Vlodavsky and Gospodarowitz, 1981; Baron-Van Evercooren *et al.*, 1982; Timpl *et al.*, 1983). Taken together, these observations suggest that laminin is a key member of a set of matrix molecules, including fibronectin, called into play during embryonic development to mediate the organization of cells into tissues and to delimit the extent and direction of cell migration.

At a molecular level, laminin interacts with a variety of other matrix components including heparin, type IV collagen, and entactin (Sakashita *et al.*, 1980; Terranova *et al.*, 1980; Carlin *et al.*, 1983). It also binds to cells, *via* a high affinity, mol. wt. 68 000, membrane protein receptor (Rao *et al.*, 1983; Brown *et al.*, 1983; Lesot *et al.*, 1983). There is evidence that the sites for these multiple interactions map to different parts of the molecule. By rotary shadowing, laminin has an asymmetric cross-like structure, with three short arms and one long arm. The short arms bear two globular regions near the ends while the long arm has a single, larger, terminal globular domain (Engel *et al.*, 1981). Proteolytic fragments which promote hepatocyte spreading appear to include at least one of the globular domains from the short arms, while the region at the end of the long arm has heparin binding activity and promotes neurite outgrowth. (Ott *et al.*, 1982; Rao *et al.*, 1983; Timpl *et al.*, 1983; Edgar *et al.*, 1984). The precise localization of the different polypeptide chains within the intact molecule is not yet known, and studies on the biosynthesis of laminin have raised important questions about its subunit composition and heterogeneity. Parietal endoderm cells of the mouse embryo produce large amounts of laminin which is assembled into a thick basement membrane (for review, see Hogan *et al.*, 1984). These cells synthesize at least three different laminin polypeptide chains with mol. wts. of ~300 000 (A chain), 205 000 (B1 chain) and 185 000 (B2 chain), all of which subsequently undergo N-glycosylation and disulphide-bonding into a molecule of mol. wt. ~900 000 (Cooper *et al.*, 1981; Kurkinen *et al.*, 1983a; Howe and Dietzschold, 1983). Recently, *in vitro* translation of mRNA from parietal endoderm cells has provided evidence for two B1 chains, migrating as a doublet on SDS-polyacrylamide gel electrophoresis (Kurkinen *et al.*, 1983a; Wang and Gudas, 1983). These chains are here termed B1a (mol. wt. 205 000) and B1b (mol. wt. 200 000). It is not known whether the different laminin subunits are assembled into molecules having the composition (A,B1a, B1b,B2), or whether heteropolymers having different subunit compositions can exist. Monoclonal antibodies recognizing laminin B chain (GP2) polypeptides show differential reaction with basement membranes from early mouse embryos and adult tissues (Wan *et al.*, 1984). This raises the possibility that there are families of laminins which differ in their structure, function and antigenicity.

Clearly, recombinant DNA techniques offer the simplest means of answering questions about the primary sequences of laminin polypeptides, the existence of a multiple gene family, and the regulation of laminin gene expression during embryonic development. Wang and Gudas (1983) have reported the isolation of a laminin B1 chain cDNA and accumulation of the corresponding mRNA in differentiating F9 cells. Here we describe the characterization and sequencing of cDNAs covering the 3′ ends of both laminin B1 and B2 mRNAs. The amino acid sequences deduced from the cDNAs provide novel information about laminin structure and predict the existence of C-terminal coiled-coil alpha-helical domains.

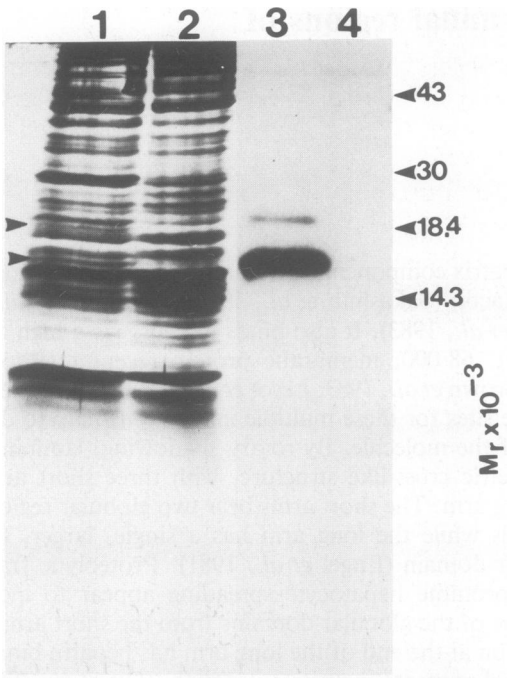

Fig. 1. Immunological identification of laminin polypeptides synthesized in *E. coli* transformed with pPE9. Extracts of log phase DH1 cells transformed with either pPE9 (**lanes 1** and **3**) or pUC8 (**lanes 2** and **4**) were analysed on a 12.5% SDS-polyacrylamide gel under reducing conditions. Proteins were either stained with silver (Morrissey, 1981) (**lanes 1** and **2**) or transferred to nitrocellulose and incubated with rabbit anti-B chain serum and [^{125}I]donkey anti-rabbit IgG, followed by autoradiography (**lanes 3** and **4**). The right hand margin shows the position of unlabelled and ^{14}C-labelled marker proteins which were included in the gel. These are ovalbumin (mol. wt. 43 000), carbonic anhydrase (30 000), β-lactoglobulin (18 400) and lysozyme (14 300).

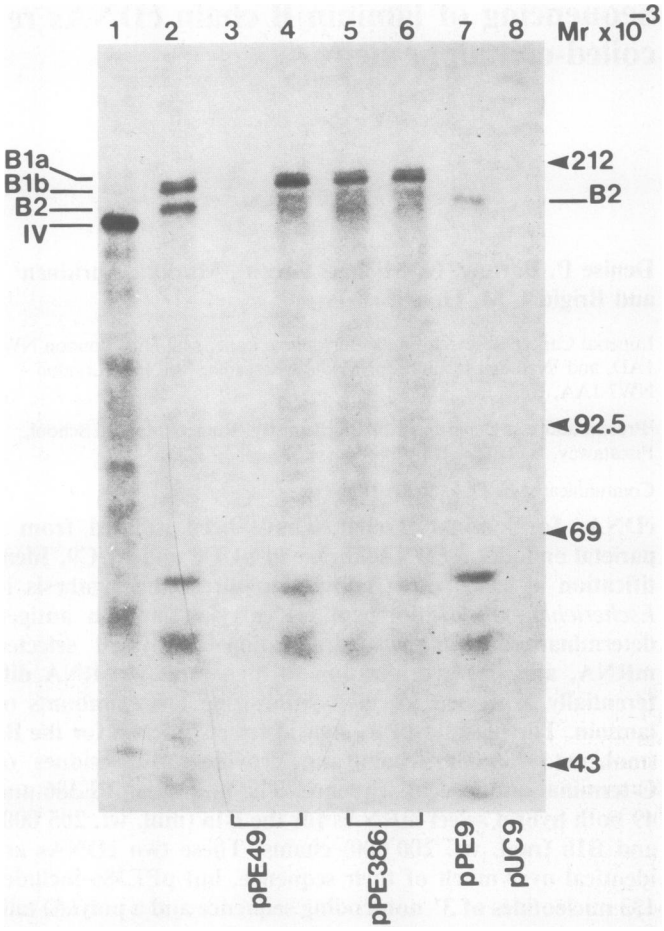

Fig. 2. Hybrid selection and *in vitro* translation. Hybrid selection was carried out using the plasmids pPE49, pPE386 and pPE9 and total RNA from the parietal endoderm cell line, F9Cl 9. mRNA eluted from the filters was translated in the nuclease-treated rabbit reticulocyte lysate. **Lane 1**: translation products of total RNA immunoprecipitated with antiserum to type IV collagen. **Lane 2**: as for **lane 1** but immunoprecipitated with anti-B chain serum. **Lanes 3** and **4**: translation products of mRNA eluted from pPE49 immunoprecipitated with anti-B chain serum. Longer exposure revealed a polypeptide doublet in **lane 3** as in **lane 4**. **Lanes 5** and **6**: translation products of mRNA eluted from pPE386 immunoprecipitated with anti-B chain serum. **Lane 7**: as **lanes 3–6** using RNA eluted from pPE9. The left hand margin shows the position of unmodified type IV collagen (mol. wt. 165 000), laminin B1a (205 000), B1b (200 000) and B2 (185 000) (Kurkinen *et al.*, 1983a). The right hand margin shows the position of the ^{14}C-labelled marker proteins (Amersham International). These are heavy chain of myosin (mol. wt. 212 000), phosphorylase B (92 500), bovine serum albumin (69 000) and ovalbumin (43 000). The immunoprecipitations were analysed on a 5–7.5% gradient SDS-polyacrylamide gel under reducing conditions. Further details of the conditions are given in Kurkinen *et al.* (1983b).

Results

Identification of a laminin B2 chain cDNA clone by immunological screening, differential Northern analysis and hybrid-select translation

Previously, we described the construction of a 20 000 member cDNA library from total parietal endoderm (PE) poly(A)$^+$ RNA in the expression vectors pUC8 and pUC9 (Kurkinen *et al.*, 1983b). The library was screened using a rabbit antiserum raised against denatured mouse laminin B chains (anti-B chain serum, Cooper *et al.*, 1981), essentially as described (Helfman *et al.*, 1983). This procedure identified one laminin fragment expressing colony, pPE9. Analysis of the proteins made by *E. coli* DH1 transformed with pPE9 revealed two unique polypeptides of mol. wt. 15 600 and 20 000. Immunoblotting showed that these polypeptides contain determinants recognized by anti-B chain serum (Figure 1).

Previous studies have shown that PE cells synthesize ~100 times more laminin than undifferentiated F9 teratocarcinoma cells, and that laminin synthesis increases 20-fold when F9 cells are treated with retinoic acid and cyclic AMP (Strickland *et al.*, 1980; Cooper *et al.*, 1983). As expected for a laminin cDNA, pPE9 hybridizes by Northern analysis to a high mol. wt. (8 kb) RNA which increases in amount when F9 cells are induced to differentiate into parietal endoderm (Figure 3 and data not shown). Evidence that pPE9 is a laminin B2 chain cDNA comes from the results of hybrid-select translation experiments (Figure 2). RNA hybridizing to pPE9 DNA immobilized on a nitrocellulose filter after elution directs the synthesis in the rabbit reticulocyte lysate of a polypeptide of mol. wt. 185 000 which is recognized by anti-B chain serum and co-migrates with B2 chains synthesized from total RNA under the same conditions.

Sequence analysis of pPE9

The 675 nucleotide insert of pPE9 was sequenced by the method of Maxam and Gilbert (1980) according to the schedule outlined in Figure 4A. It contains a single open reading frame, extending to position 652 and encoding 217 amino acids, followed by a stop codon and 20 nucleotides of 3′ non-coding sequence (Figure 5A). As shown in Figure 1, bacteria transformed with pPE9 make polypeptides containing laminin antigenic determinants. However, the reading

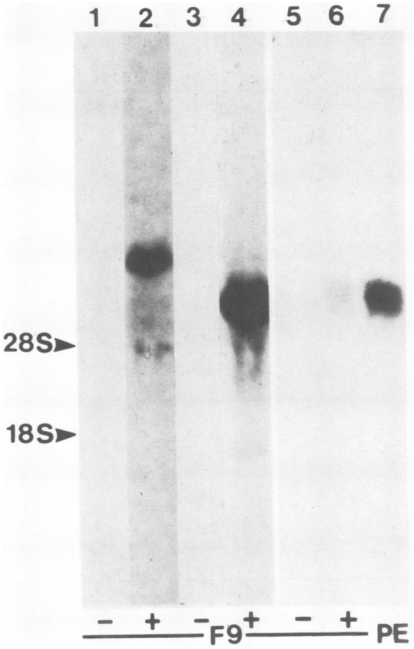

Fig. 3. Differential Northern analysis. Total RNA from undifferentiated (F9 −) and differentiated (F9 +) F9 cells and from mouse embryo PE cells was separated on a 0.7% formaldehyde agarose gel, transferred to nitrocellulose and probed with nick-translated plasmids as described (Kurkinen *et al.*, 1983b) **Lanes 1** and **2**: pPE9 hybridizes to an 8-kb RNA differentially expressed in F9 + cells. **Lanes 3** and **4**: pPE49 hybridizes to a 6-kb RNA differentially expressed in F9 + cells. **Lanes 5 – 7**: pPE386 hybridizes to a 6-kb RNA expressed in F9 + cells and at higher levels in PE cells. Full details of the kinetics of induction of the RNAs in these experiments will be described elsewhere (D.P. Barlow, in preparation). The left hand margin shows the position of the 28S and 18S rRNA in the gel. These were used to calculate the sizes of the RNA.

frame initiating at the AUG of the β-galactosidase gene in the pUC8 vector (Helfman *et al.*, 1983) terminates with a stop codon after nucleotide 21 of the cDNA insert, excluding the possibility of readthrough into a fusion protein. It is therefore likely that internal initiation occurs within the insert, as reported in similar systems by others (Kemp and Cowman, 1981). Initiation at the AUG at position 215, and at the cluster of three AUGs near position 281, would generate proteins with predicted mol. wts. of 13 726 and 16 321, which are close to the size of the two polypeptides observed by immunoblotting (Figure 1).

Identification of laminin B1 chain cDNAs by differential hybridization, Northern analysis, and hybrid-select translation

Elsewhere, we have described the isolation, from a random subset of the pUC9 PE library, of 79 plasmids differentially hybridizing to a cDNA probe from PE cells compared with undifferentiated F9 cells (Kurkinen *et al.*, 1983b). Sixty of these clones were further characterized by Northern analysis. Just one of the clones, pPE386, hybridized to a 6-kb RNA present at high levels in PE and differentiated F9 cells, but absent from undifferentiated F9 cells (Figure 3). As shown in Figure 2, RNA hybrid selected by pPE386 directs the *in vitro* synthesis of two polypeptides of mol. wt. 205 000 and 200 000. These chains are recognized by anti-B chain serum and co-migrate on SDS-polyacrylamide gels with the B1a and B1b chains synthesized in the same lysate supplemented with total RNA from parietal endoderm cells. Plasmid pPE386, which contains a 1.1-kb insert, was then used to screen the complete PE library in both pUC8 and pUC9. Surprisingly, this resulted in the isolation of only one other plasmid, pPE49. This cDNA also hybrid-selects mRNAs which direct the synthesis of B1a and B1b polypeptides in the reticulocyte lysate (Figure 2).

Sequence analysis of pPE49 and 386

The inserts of these two plasmids were sequenced by a combination of Maxam and Gilbert and M13 dideoxy methods according to the strategies outlined in Figure 4B. The 1120 nucleotide sequence of pPE49 gives a single open reading frame encoding 372 amino acids, or ~20% of a laminin B1 chain (Figure 5B).

pPE386 has an identical sequence to pPE49 over much of its length but extends in the 3′ direction for a further 123 amino acids, before terminating at position 993. The stop codon is followed by 133 bases of non-coding sequence, a poly(A) addition site, and a poly(A) tail. At the 5′ end, the sequences of the two cDNAs differ up to nucleotide 116 of pPE386. Continuity of the reading frame through nucleotides 1 – 116 is prevented by two stop codons at positions 51 and 66. It is therefore likely that the 5′ unique sequence of pPE386 either incorporates a cloning artifact or is derived from an unprocessed nuclear RNA sequence (see Discussion). For the purposes of analysing the predicted amino acid sequence below we have assumed that pPE49 and pPE386 are overlapping cDNAs for the same B1 chain, and that there is continuity from the 3′ end of the reading frame of pPE49 (see arrow in Figure 5B) into the same reading frame of pPE386.

Analysis of predicted amino acid sequence of B2 fragment pPE9

The nucleotide sequence of pPE9 predicts a polypeptide fragment of 217 amino acids, finishing at the C terminus of the B2 subunit. Analysis of the secondary structure by a programme based on the method of Chou and Fasman (Roberts and Geisow, 1980) showed that this fragment was predominantly alpha-helical. Closer inspection of the sequence, written in the form of an alpha-helical net, revealed a remarkable density of hydrophobic residues at positions *a* (90%) and *d* (70%) of a repeating seven residue segment (*a b c d e f g*).[If the cylindrical surface of an alpha helix is cut and opened out onto a plane, the amino acid residues form a net which repeats every 3.6 residues. To a first approximation a heptad of amino acids (*a b c d e f g*) forms two turns. Residues in positions *a* and *d* form a band down one edge of the helix which twists slowly round it. If these *a* and *d* residues are hydrophobic, a pair of similar helices can coil around each other, burying the hydrophobic residues and giving a stable coiled-coil, parallel double helix. This was first envisaged by Crick (1953) and later observed in tropomyosin (Sodek *et al.*, 1972) and many other proteins.] This produces a characteristic hydrophobic band down one side of a long helix as previously observed in tropomyosin (Sodek *et al.*, 1972; McLachlan and Stewart, 1976), light meromyosin (McLachlan and Karn, 1983) and intermediate filament proteins (Geisler and Weber, 1982) and fibrinogen (Doolittle *et al.*, 1978). This immediately suggests that the C terminus of laminin B2 chain participates in a coiled-coil two or three stranded helix typical of these proteins. The regularity of the repeats in the B2 chain fragment is emphasized in Figure 6A where the sequence is aligned in 28 residue segments, interrupted by two insertions.

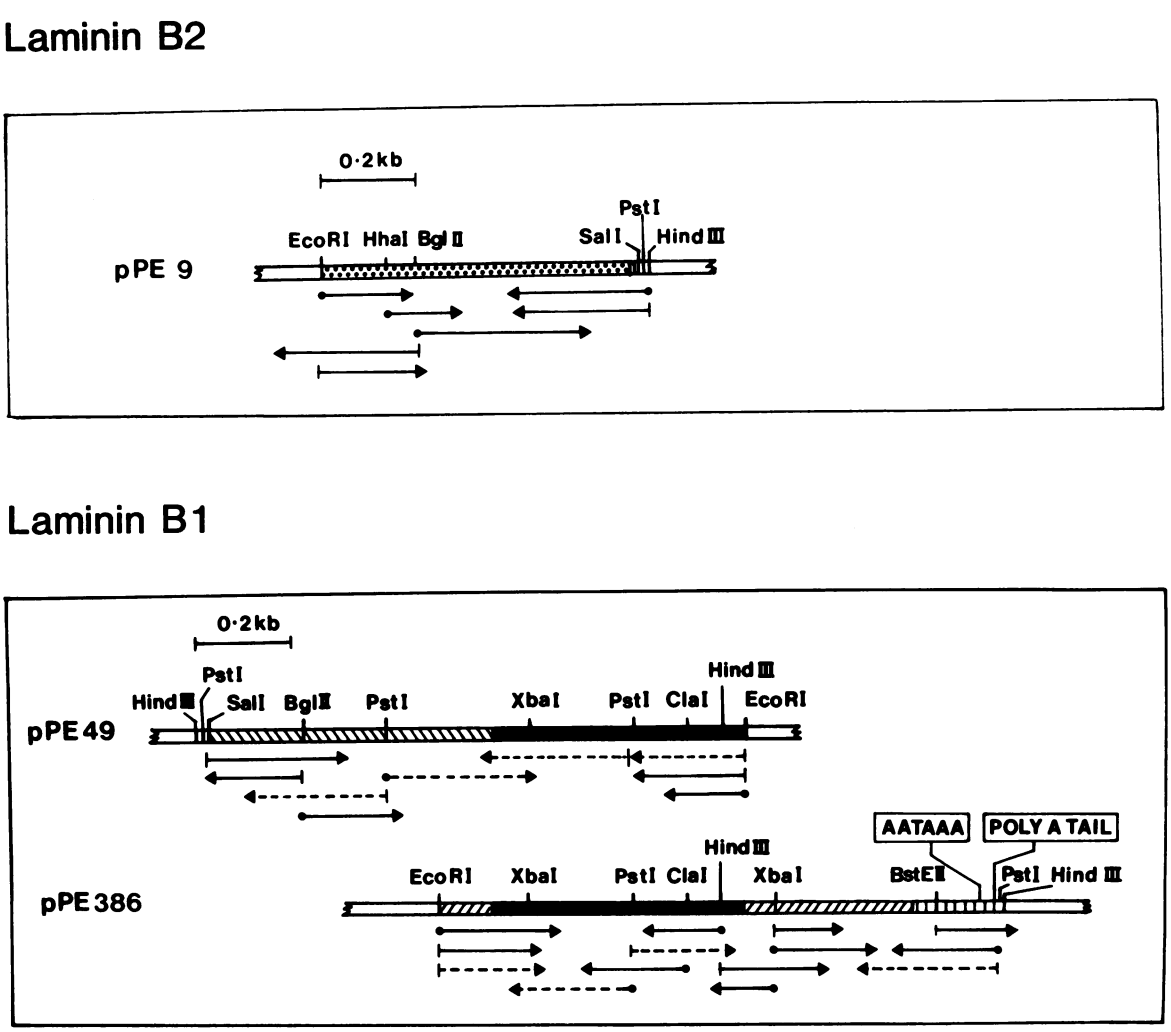

Fig. 4. Sequencing strategy and restriction maps for B1 and B2 laminin cDNA chains. The cDNA was sequenced either by chemical cleavage (——▶) or by dideoxy chain termination (- - -▶) of the sense (●——) or anti-sense (|——) strand. Vector DNA is unshaded. Sequences unique to B2 pPE9 are dotted. Sequences identical in pPE49 and pPE386 are solid, while sequences unique to either pPE49 or pPE386 are cross-hatched. The 3′ non-coding sequences of pPE386 are marked with vertical lines.

Analysis of predicted amino acid sequences of B1 fragment pPE49 and 386

The amino acid sequence deduced from the combined nucleotide sequences of pPE49 and 386 consists of 495 residues from the C terminus of a B1 subunit. The method of Chou and Fasman again predicts mainly alpha helix. The C-terminal 354 residues show the same remarkable regularity that was found in the B2 fragment and this section of the sequence is shown in Figure 6B, written in a similar format. It is clear from this regular distribution of hydrophobic residues that the C termini of both B1 and B2 could participate in extended coiled-coil double or triple helices, with occasional breaks. The 354 residues of B1 could give ~50 nm of helix in partnership with another subunit(s). The other striking feature of the repeats illustrated in Figure 6A and B is the high concentration of charged residues (~50%) on sites *b c e f* and *g*. This pattern is characteristic of tropomyosin and the tail of myosin, the charge density being much higher than that seen in keratin and desmin (Geisler and Weber, 1982). An analysis of the preference of the different sites of the heptad for charged and hydrophobic residues is given in Table I, which shows a fairly uniform distribution of both positive and negative charge over all the non-hydrophobic sites of B1 and B2. There is no complementarity of charge between sites *e* and *g* as was observed in tropomyosin (McLachlan and Stewart, 1976). Some evidence for a longer range periodicity of charge can be seen in Figure 6A and B, which emphasizes a high density of negative charge on the first half of the second heptad, followed immediately by a concentration of positively charged residues. There are further weaker alternations of charge bands along the helix axis which may well have a significance for the interactions of laminin subunits with each other and with different molecules, analogous to that of the charge patterns in myosin (McLachlan and Karn, 1983) and tropomyosin (McLachlan and Stewart, 1976). A more detailed analysis of the periodicity will be left until more complete sequences are available.

The N-terminal 106 residues of B2 are also strongly predicted as mainly alpha-helical, divided into three or four short helices separated by bends and lacking the heptad repeats. They are separated from the repetitive region by 35 residues which include 10 glycines, four prolines and six cysteines. This sequence shows some similarity to that found in the cysteine-rich segment of fibrinogen which is involved in inter-helix cross-linking (Doolittle *et al.*, 1978).

Laminin B2

```
                PPE9 1
ATGATTACGAATTCC C GCC ATC AAC CGG ACC ATA GCT GAA GCC AAT GAG AAG ACA AGG GAG GCC CAG CTA GCG CTG GGC AAT GCT GCC GCT GAC GCC ACG GAG GCC AAG  -  94
                A  I  N* R  T  I  A  E  A  N  E  K  T  R  E  A  Q  L  A  L  G  N  A  A  A  D  A  T  E  A  K   -  31

AAC AAG GCC CAT GAG GCA GAG AGG ATT GCC AGC GCC GTG CAG AAG AAT GCC ACC AGT ACC AAG GCG GAC GCA GAA AGA ACC TTC GGG GAA GTT ACA GAT CTG GAT AAT  - 202
N  K  A  H  E  A  E  R  I  A  S  A  V  Q  K  N* A  T  S  T  K  A  D  A  E  R  T  F  G  E  V  T  D  L  D  N   -  67

GAG GTG AAT GGT ATG CTG AGG CAG CTG GAG GAG GCA GAG AAT GAG CTG AAG AAG AAG CAA GAT GAT GCC GAC CAG GAC ATG ATG ATG GCA GGG ATG GCT TCG CAG GCT  - 310
E  V  N  G  M  L  R  Q  L  E  E  A  E  N  E  L  K  K  K  Q  D  D  A  D  Q  D  M  M  M  A  G  M  A  S  Q  A   - 103

GCT CAG GAG GCT GAG CTC AAT GCC AGA AAG GCC AAA AAC TCT GTC AGC AGC CTC CTC AGC CAG CTG AAC AAC CTC TTG GAT CAG CTA GGA CAG CTG GAC ACA GTG GAC  - 418
A  Q  E  A  E  L  N  A  R  K  A  K  N  S  V  S  S  L  L  S  Q  L  N  N  L  L  D  Q  L  G  Q  L  D  T  V  D   - 139

CTG AAC AAG CTC AAT GAG ATC GAA GGC TCC CTG AAC AAA GCC AAA GAC GAA ATG AAG GCC AGC GAC CTG GAC AGG AAG GTG TCT GAC CTG GAG AGC GAG GCT CGG AAG  - 526
L  N  K  L  N  E  I  E  G  S  L  N  K  A  K  D  E  M  K  A  S  D  L  D  R  K  V  S  D  L  E  S  E  A  R  K   - 175

CAG GAG GCA GCC ATC ATG GAC TAT AAC CGG AAC ATA GCA GAG ATC ATT AAG GAT ATT CAC AAC CTG GAG GAC ATC AAG AAG ACC CTA CCA ACC GGC TGC TTC AAC ACC  - 634
Q  E  A  A  I  M  D  Y  N  R  N  I  A  E  I  I  K  D  I  H  N  L  E  D  I  K  K  T  L  P  T  G  C  F  N  T   - 211

CCG TCC ATC GAG AAG CCC TAG TGGCGAGAGG GCTGTAAGGG  - 675
P  S  I  E  K  P  •                                - 217
```

Laminin B1

```
 CA GCA GAG AGC CTT GAC AAG ACC GTG AAG GAG CTG GCA GAA CAG CTG GAG TTT ATC AAA AAC TCC GAT ATT CAG GGC GCC TTG GAT AGC ATC ACC AAG TAT TTC CAG  - 107
    A  E  S  L  D  K  T  V  K  E  L  A  E  Q  L  E  F  I  K  N  S  D  I  Q  G  A  L  D  S  I  T  K  Y  F  Q   -  35

ATG TCT CTT GAG GCA GAG AAG CGG GTG AAT GCC TCC ACC ACA GAC CCC AAC AGC ACT GTG GAG CAG TCT GCC CTC ACG CGA GAC AGA GTA GAA GAT CTG ATG TTG GAG  - 215
M  S  L  E  A  E  K  R  V  N* A  S  T  T  D  P  N  S  T  V  E  Q  S  A  L  T  R  D  R  V  E  D  L  M  L  E   -  71

CGA GAG TCT CCG TTC AAG GAG CAG CAG GAG GAA CAG GCA CGC CTC CTG GAC GAA CTG GCC GGC AAA CTG CAA AGT CTC GAC CTG TCG GCT GCT GCA CAG ATG ACC TGT  - 323
R  E  S  P  F  K  E  Q  Q  E  E  Q  A  R  L  L  D  E  L  A  G  K  L  Q  S  L  D  L  S  A  A  A  Q  M  T  C   - 107

GGA ACA CCT CCA GGG GCT GAC TGT TCT GAA AGT GAA TGT GGT GGC CCC AAC TGC AGA ACT GAC GAA GGA GAG AAG AAG TGT GGG GGG CCT GGC TGT GGT GGT CTG GTC  - 431
G  T  P  P  G  A  D  C  S  E  S  E  C  G  G  P  N  C  R  T  D  E  G  E  K  K  C  G  G  P  G  C  G  G  L  V   - 143

                                                                  pPE386  -  CC GCC CAA AGA GGT CTA CCT GGA CAG GAG CCT GCT GAC CCT GGA
ACT GTG GCC CAC AGT GCT TGG CAG AAA GCC ATG GAT TTT GAC CGT GAT GTC CTG AGT GCC CTG GCT GAA GTC GAA CAG CTC TCC AAG ATG GTC TCT GAA GCA AAA GTG  - 539
T  V  A  H  S  A  W  Q  K  A  M  D  F  D  R  D  V  L  S  A  L  A  E  V  E  Q  L  S  K  M  V  S  E  A  K  V   - 179

GGA CAA TGA ACT GGG CTC CGG TAA CTT CGG GAC TGT GAA AAA GGG ATA CTA CCA AAT GAA AAA AGT TGT GAA AAG AGC --- --- --- --->
AGA GCA GAT GAG GCG AAG CAG AAT GCG CAG GAT GTC CTG TTA AAA ACA AAT GCT ACC AAA GAA AAA GTG GAC AAG AGC AAC GAG GAC CTG CGG AAC CTC ATC AAG CAG  - 647
R  A  D  E  A  K  Q  N  A  Q  D  V  L  L  K  T  N* A  T  K  E  K  V  D  K  S  N  E  D  L  R  N  L  I  K  Q   - 215

ATC AGA AAC TTC CTG ACT GAG GAT AGT GCT GAT CTA GAC AGT ATT GAA GCA GTT GCT AAT GAA GTA CTG AAA ATG GAA ATG CCT AGC ACG CCA CAG CAG TTA CAG AAC  - 755
I  R  N  F  L  T  E  D  S  A  D  L  D  S  I  E  A  V  A  N  E  V  L  K  M  E  M  P  S  T  P  Q  Q  L  Q  N*  - 251

CTA ACA GAA GAC ATT CGG GAG CGA GTT GAA ACC CTC TCT CAA GTA GAG GTT ATT TTG CAG CAG AGT GCA GCT GAC ATT GCC AGA GCT GAG CTG TTG CTT GAG GAA GCT  - 863
L  T  E  D  I  R  E  R  V  E  T  L  S  Q  V  E  V  I  L  Q  Q  S  A  A  D  I  A  R  A  E  L  L  L  E  E  A   - 287

AAG AGA GCA AGC AAA AGT GCA ACA GAT GTT AAA GTC ACT GCA GAC ATG GTG AAG GAA GCA TTA GAA GAA GCA GAA AAG GCC CAG GTT GCA GCA GAG AAG GCG ATT AAA  - 971
K  R  A  S  K  S  A  T  D  V  K  V  T  A  D  M  V  K  E  A  L  E  E  A  E  K  A  Q  V  A  A  E  K  A  I  K   - 323

CAA GCT GAT GAG GAT ATC CAA GGA ACC CAA AAC CTG CTA ACA TCG ATT GAA TCT GAA ACG GCA GCT TCT GAG GAA ACC CTG ACC AAC GCC TCC CAG CGC ATC AGC AAG -1079
Q  A  D  E  D  I  Q  G  T  Q  N  L  L  T  S  I  E  S  E  T  A  A  S  E  E  T  L  T  N* A  S  Q  R  I  S  K   - 359

CTT GAG AGG AAC GTG GAA GAG CTT AAG CGT AAA GCT GCC CAG AAC TCT GGG GAG GCA GAA TAT ATC GAA AAA GTA GTA TAT TCT GTA AAA CAG AAT GCA GAC GAT GTT -1187
L  E  R  N  V  E  E  L  K  R  K  A  A  Q  N  S  G  E  A  E  Y  I  E  K  V  V  Y  S  V  K  Q  N  A  D  D  V   - 395

AAG AAG ACT CTA GAT TGC GAA CTT GAT GAA AAG TAT AAG AAG GTA GAA AGT TTA ATT GCC CAA AAA ACT GAA GAG TCA GCA GAT GCC AGG AGG AAA GCT GAG CTG CTA -1295
K  K  T  L  D  C  E  L  D  E  K  Y  K  K  V  E  S  L  I  A  Q  K  T  E  E  S  A  D  A  R  R  K  A  E  L  L   - 431

CAA AAT GAA GCA AAA ACA CTC TTG GCT CAA GCT AAC AGC AAG CTC CAG CTG TTG GAA GAC TTA GAA AGA AAA TAT GAG AAC AAT CAA AAA TAC TTA GAA GAT AAA GCT -1403
Q  N  E  A  K  T  L  L  A  Q  A  N  S  K  L  Q  L  L  E  D  L  E  R  K  Y  E  N  N  Q  K  Y  L  E  D  K  A   - 467

CAA GAA TTG GTG CGA CTG GAA GGA GAG GTT CGC TCC CTC CTT AAG GAC ATA AGT GAG AAA GTT GCG GTT TAC AGC ACC TGC TTA TAA CAGGAAGGGG CTGTAGAGGG GCTCG -1517
Q  E  L  V  R  L  E  G  E  V  R  S  L  L  K  D  I  S  E  K  V  A  V  Y  S  T  C  L  •                          - 495

GTGAC CAAGGTAAAC CACACGCGCA AACCGAGGCA GTCATCTACA AATAACCCAT CATCTATTTA ATGTTTTTAA CCACCTACTT TTGTATGGAG TTAAATAAAA GACATTGGTT TTGTATAAAC A(A18) -1654
```

Fig. 5. DNA sequence of laminin cDNAs. Laminin B2 – pPE9. Underlined are pUC8 vector sequences and potential sites for internal initiation of protein synthesis on RNA transcribed in *E. coli*. The predicted amino acid sequence for the open reading frame extending to nucleotide 653 is shown below the DNA sequence using the one letter code. Potential asparagine glycosylation sites are marked with an asterisk, and the single cysteine with a circle. Laminin B1 – pPE49 and 386. The DNA sequence and predicted amino acid sequence covered by the two overlapping cDNAs are shown. The most 5′ sequences up to the arrow are from pPE49 and the most 3′ sequences from pPE386. The most 5′ 116 bp of pPE386 which differ from the corresponding overlapping region of pPE49 are shown in italics. These 116 bp are either an intron sequence or a cloning artifact and are not included in subsequent analysis (see text). The polyadenylation signal and the poly(A) tail are underlined. Potential glycosylation sites and cysteine residues are marked as in pPE9. Note that six of the cysteines are clustered together.

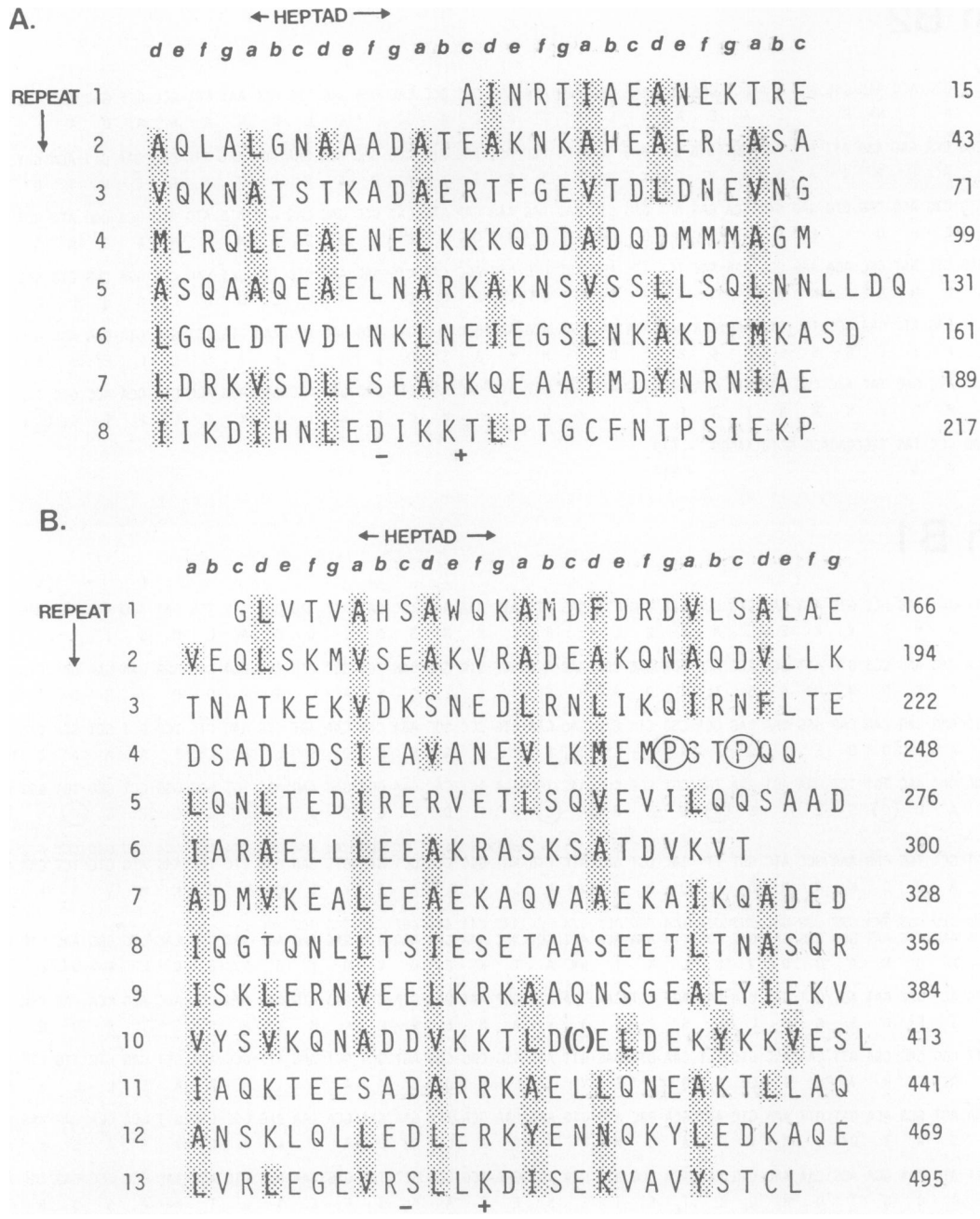

Fig. 6. Periodicity of the C-terminal fragments of laminin B2 and B1 subunits. The sequences are written as repeats of 28 residues (covering four heptads) and labelled according to the scheme used for tropomyosin (McLachlan and Stewart, 1976). **(A)** The B2 fragment shows two interruptions to the hydrophobic periodicity. The first is an insertion of an extra *d, e, f, g* tetrad in the fifth repeat which disturbs the regular alternation of tetrads and triads but would distort the helix only slightly. The second is an addition of two residues after the sixth repeat which shifts the hydrophobic core around the helix by 180° and implies a break in the regular structure. There are, in addition, three prolines in the last 13 residues but these are outside the repeating region. **(B)** The B1 fragment (residues 142−495) which shows only one major discontinuity produced by a nine residue segment containing two prolines (no. 243 and 246, ringed) followed by a change in phase in the alternation of tetrads and triads. An extra residue **(C)** is inserted at position 401 but this could be accommodated without major distortion. In both sequences negative charges are concentrated in the first half of the second heptad of each 28 residue repeat (−), while positive charges are abundant in the second half (+).

Discussion

Isolation of laminin cDNAs from the PE library

Screening of the PE cDNA library by two different methods has resulted in the isolation of three cDNAs for laminin B chains identified on the basis of the following criteria: (i) differential hybridization to cDNA from PE cells (Kurkinen *et al.*, 1983b), (ii) hybridization to high mol. wt. RNA which is abundant in PE and which increases when F9 cells are induced to differentiate, (iii) hybrid selection of mRNAs which direct the *in vitro* synthesis of laminin B chain polypeptides, and (iv), in the case of pPE9, synthesis in *E. coli* of fragments carrying laminin B chain determinants. PE cells synthesize large amounts of laminin, and we have calculated that ~0.4% of the total mRNA of these cells codes for individual laminin polypeptide chains (Hogan *et al.*, 1984). It was therefore surprising that screening of the complete library with the insert of pPE386 led to the isolation of only one other laminin

Table I. Distribution of charged and hydrophobic residues in the repetitive sections of B1 and B2

Site class	Site	Hydrophobic[a] residues/ 100 sites		Negative charge/ 100 sites		Positive charge/ 100 sites	
		B_1	B_2	B_1	B_2	B_1	B_2
Hydrophobic core	*a*	92	87	2	3	0	3
	d	76	73	2	3	8	3
Core boundary	*e*	32	23	28	33	18	13
	g	20	23	30	33	14	10
	b	22	7	30	10	16	23
External	*c*	16	13	28	40	14	17
	f	8	20	22	17	28	23

The seven sites of each heptad are subdivided into hydrophobic core (*a,d*), core boundary (*e,g*) and external sites (*b,c,f*).
[a]Ala, Val, Leu, Ile, Met, Phe, Tyr, Trp are regarded as hydrophobic.

B1 chain cDNA. In contrast, screening under the same conditions with the B2 chain cDNA, pPE9, and several other PE-specific cDNAs including ones for type IV collagen, has led to the isolation of between 30 and 300 related cDNAs (unpublished results). At present we have no explanation for the underrepresentation of laminin B1 chain cDNAs in our library.

Both pPE386 and pPE49 hybrid-select RNAs which direct the *in vitro* synthesis of laminin B1a and B1b polypeptides. Similar results have been observed for the B1 chain cDNA described by Wang and Gudas (1983). This strongly suggests that B1a and B1b are coded for by closely related mRNAs, possibly derived from the same gene by alternate splicing. pPE386 differs from pPE49 over the first 116 nucleotides. However, the absence of a continuous reading frame in this region suggests that the most 5′ unique sequence of pPE386 is derived either from a cloning artifact or from the unprocessed nuclear RNA precursor. However, until more cDNAs can be analysed we cannot clearly distinguish between these and other explanations.

Comparison of the sequences of pPE9, 386 and 49 shows that there is no significant homology between the laminin B1 and B2 chains at the nucleotide level. It is also interesting that the two polypeptides are coded for by mRNAs of very different size. pPE386 and 49 both hybridize to a 6-kb RNA. This is close to the minimum size for a mRNA coding for a protein of mol. wt. ~200 000, and is in good agreement with the presence of only 133 nucleotides of 3′ non-coding sequence in pPE386. pPE9, on the other hand, hybridizes to an 8-kb RNA, suggesting that the B2 chain mRNA which codes for a protein of mol. wt. 185 000 has ~2.5 kb of non-coding sequence.

Location of predicted structures within the laminin molecule

The work of Engel *et al.* (1981) and Timpl *et al.* (1983) has established a cross-like morphology for the intact laminin molecule (mol. wt. 900 000), and their study of proteolytic fragments has characterized the globular domains located at the ends of the four arms of the cross in terms of both amino acid composition and immunological and substrate binding properties. However, the exact arrangement of the different polypeptide chains in the intact molecule is not at all clear. It is probable that they are covalently linked by the 50 or 60 disulphide bonds found in the central protease resistant part of the three shorter arms of the cross (proteolytic fragments P1, E1 and S1, Rohde *et al.*, 1980; Engel *et al.*, 1981; Ott *et al.*, 1982).

From the point of view of establishing the location of the alpha-helical regions predicted for our B1 and B2 chain fragments, the circular dichroism measurements of Ott *et al.* (1982) are of great interest. About 30% of the laminin molecule, or ~2500 amino acid residues, are in alpha-helical configuration, and about half of them unfold when the protein is heated, in a remarkably sharp transition between 57°C and 61°C. The helices in tropomyosin and the myosin tail melt somewhat less sharply in a series of transitions between 40°C and 60°C (Privalov, 1982). These two proteins and the helical regions of laminin are also unfolded at room temperature by 2 M guanidinium chloride (Lehrer, 1978; Privalov, 1982; Ott *et al.* 1982). These parallels in stability confirm the close resemblance of the helical regions of laminin to those of myosin and tropomyosin. Quantitatively the same loss of alpha-helix from native laminin is observed when it is digested with elastase (37°C, 24 h, Ott *et al.*, 1982) as when it is heated to 60°C, and the resulting elastase-resistant fragments have little alpha-helical circular dichroism. The pepsin-resistant fragment, P1, also lacks alpha-helical structure. These observations lead to the conclusion that the temperature-sensitive alpha-helical regions do not survive the proteolysis used to produce fragments P1, E1 and S1. This suggestion is supported by a comparison of the amino acid composition of the complete laminin molecule with that of the sum of the protease-resistant fragments (Ott *et al.*, 1982; Timpl *et al.*, 1983). This shows a large deficiency in glutamic, aspartic, alanine and leucine residues (the most abundant residues of the helical sections of B1 and B2) while the glycines, prolines, phenylalanines, histidines and tyrosines (which make up only 4% of the helical regions) are well accounted for. Since electron microscopic characterization of the protease fragment P1 shows that it represents the central 20 nm of each of the three short arms of the cross, the labile helical regions must be either in the 10 nm segments between the paired globular domains of the short arms or, more likely, in the long arm. If this localization is correct then the C-terminal sections of B1 and B2 should also form part of the long arm.

The question of whether laminin contains double or triple helices remains open. In general, double helices are more common, but the triple helix of fibrinogen is well established, although it is less regular than that of the myosins. The mass per unit length of the arms of laminin (Engel *et al.*, 1981) is close to that expected for triplet or paired double helices and such possibilities must be borne in mind. Clearly a complete answer to the question of the location and heterogeneity of laminin B chains in the intact molecule requires in part the isolation and sequencing of longer cDNAs. It may also be possible to map by rotary shadowing techniques the location of those amino acid sequences already deduced using antibodies raised against fragments made in bacteria (see Figure 1) or against synthetic peptides.

Materials and methods

The sources of most reagents have been described (Kurkinen *et al.*, 1983a,1983b).

Antisera

The rabbit anti-mouse laminin B chain serum has been described (Cooper *et*

al., 1981). Rabbit anti-mouse type IV collagen serum was a gift from Dr. E.D. Adamson.

Hybrid-select translation

This was carried out as described (Maniatis *et al.*, 1982; Kurkinen *et al.*, 1983b) using 25 μg plasmid DNA bound to 3 mm^2 nitrocellulose filters (Sartorius) and 1 mg/ml total RNA isolated from the parietal endoderm cell line, F9Cl 9.

RNA isolation and Northern analysis

Total RNA was isolated by a modification of the guanidinium thiocyanate CsCl centrifugation technique as described (Kurkinen *et al.*, 1983a). 10 μg total cell RNA was analysed on 0.7% formaldehyde agarose gels and transferred to nitrocellulose (Sartorius) for hybridization to ^{32}P-labelled, nick-translated plasmids. All conditions were as described (Kurkinen *et al.*, 1983b). Ethidium bromide fluorescence was used to locate the position of the rRNA, and to confirm that equal amounts had been loaded.

DNA sequencing

For sequencing by chemical cleavage (Maxam and Gilbert, 1980) cDNA fragments with protruding ends were labelled either at the 5′ end of the strand using T4 polynucleotide kinase (Weaver and Weissman, 1979) and γ-^{32}P-labelled deoxyribonucleoside triphosphate, or at the 3′ end of the strand using Klenow fragment and α-^{32}P-labelled deoxyribonucleoside triphosphate. Labelled fragments were separated on low melting agarose and excised. The agarose was melted at 70°C, phenol extracted, made 1.24 M with ammonium acetate and ethanol precipitated. The precipitate was taken up in water and lyophilised. The method also included a modification of the piperidine step as described by Smith and Calvo (1980). For sequencing by the Sanger dideoxy chain terminating technique (Sanger *et al.*, 1977) appropriate cDNA restriction fragments were subcloned into M13 mp8 and mp9. Preparation and sequencing of the single-stranded template was performed using the instructions and kit supplied by Amersham International plc.

Cell culture

The parietal endoderm cell line, F9Cl 9 (Solter *et al.*, 1979) was obtained from Dr. C. Graham. F9 teratocarcinoma cells were grown in Dulbecco's modified Eagle's medium containing 10% foetal bovine serum on tissue culture dishes pre-treated with 1% gelatin. For isolation of RNA, stock cultures were seeded at a density of ~1 x $10^4/cm^2$ into 144 cm^2 culture dishes.

Control cells were harvested after 15 h. At the same time differentiation was initiated by the addition of 5 x 10^{-8} M retinoic acid (Sigma all trans) 0.1 mM dibutyryl cyclic AMP (Sigma) and 0.1 mM isobutyl methyl xanthine (Sigma). Differentiated cells were harvested after 4 days.

Immunoblotting

200 μl of log phase *E. coli* DH1 cells were washed twice in phosphate-buffered saline and resuspended in the same volume of 0.3 M KCl, 0.01 M EDTA, 0.01 M dithiothreitol, 0.0005 M phenylmethylsulphonylfluoride and 0.05 M Tris, pH 7.5. The suspension was sonicated on ice (three 5 s bursts of 14 μm amplitude) and cleared by centrifugation at 10 000 *g* for 10 min. Aliquots of the supernatant containing equal amounts of protein were analysed on a 12.5% SDS-polyacrylamide gel under reducing conditions (Laemmli, 1970). Proteins were transferred to nitrocellulose and incubated with a 1:200 dilution of rabbit anti-laminin B chain serum followed by 0.4 μCi donkey anti-rabbit IgG (Amersham International, 10 μCi/μg) as described (Hogan *et al.*, 1982).

Acknowledgements

We thank Dr. Ian Teo for advice with the Western blotting, Drs. Colin Hughes and Michael Geisow for helpful discussions, and Peter Goldsbrough and Amanda Taylor for technical assistance.

References

Baron-Van Evercooren,A., Kleinman,H.K., Ohno,S., Marangos,P., Schwart,J.P. and Dubois-Dalcq,M.E. (1982) *J. Neurosci. Res.*, **8**, 179-193.
Brown,S.S., Malinoff,H.L. and Wicha,M.S. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 5927-5930.
Carlin,B.E., Durkin,M.E., Bender,B., Jaffe,R. and Chung,A.E. (1983) *J. Biol. Chem.*, **258**, 7729-7737.
Cooper,A.R. and MacQueen,H. (1983) *Dev. Biol.*, **96**, 467-471.
Cooper,A.R., Kurkinen,M., Taylor,A. and Hogan,B.L.M. (1981) *Eur. J. Biochem.*, **119**, 189-197.
Cooper,A.R., Taylor,A. and Hogan,B.L.M. (1983) *Dev. Biol.*, **99**, 510-516.
Crick,F.H.C. (1953) *Acta Crystallogr.*, **6**, 689-697.
Doolittle,R.F., Goldbaum,D.M. and Doolittle,L.R. (1978) *J. Mol. Biol.*, **120**, 311-325.
Edgar,D., Timpl,R. and Thoenen,H. (1984) *EMBO J.*, **3**, 1463-1468.
Engel,J., Odermatt,E., Engel,A., Madri,J.A., Furthmayr,H., Rhode,H. and Timpl,R. (1981) *J. Mol. Biol.*, **150**, 97-120.
Foidart,J.-M., Bere,E.W., Yaar,M., Rennard,S.I., Gullino,M., Martin,G.R. and Katz,S.I. (1980) *Lab. Invest.*, **42**, 336-342.
Geisler,N. and Weber,K. (1982) *EMBO J.*, **1**, 1649-1656.
Helfman,D.M., Feramisco,J.R., Fiddes,J.C., Thomas,G.P. and Hughes, S.H. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 31-35.
Hogan,B.L.M., Taylor,A., Kurkinen,M. and Couchman,J.R. (1982) *J. Cell Biol.*, **95**, 197-204.
Hogan,B.L.M., Barlow,D.P. and Kurkinen,M. (1984) in *CIBA Symposium on Basement Membranes and Cell Movement*, in press.
Howe,C.C. and Dietzschold,B. (1983) *Dev. Biol.*, **98**, 385-391.
Kemp,D.J. and Cowman,A.F. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 4520-4524.
Kuhl,U., Timpl,R. and von der Mark,K. (1982) *Dev. Biol.*, **93**, 344-354.
Kurkinen,M., Barlow,D.P., Jenkins,J.R. and Hogan,B.L.M. (1983a) *J. Biol. Chem.*, **258**, 6543-6548.
Kurkinen,M., Barlow,D.P., Helfman,D.M., Williams,J.R. and Hogan, B.L.M. (1983b) *Nucleic Acids Res.*, **11**, 6199-6209.
Laemmli,U.K. (1970) *Nature*, **227**, 680-685.
Lehrer,S.S. (1978) *J. Mol. Biol.*, **118**, 209-226.
Leivo,I., Vaheri,A., Timpl,R. and Wartiovaara,J. (1980) *Dev. Biol.*, **76**, 100-114.
Lesot,H., Kuhl,U. and von der Mark,K. (1983) *EMBO J.*, **2**, 861-868.
Liesi,P., Dahl,D. and Vaheri,A. (1983) *J. Cell Biol.*, **96**, 920-924.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*, published by Cold Spring Harbor Laboratory Press, NY.
Martinez-Hernandez,A. and Chung,A.E. (1984) *J. Histochem.*, **32**, 289-298.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499-560.
McLachlan,A.D. and Stewart,M. (1976) *J. Mol. Biol.*, **103**, 271-298.
McLachlan,A.D. and Karn,J. (1983) *J. Mol. Biol.*, **164**, 605-626.
Morrissey,J.H. (1981) *Anal. Biochem.*, **117**, 307-310.
Ott,U., Odermatt,E., Engel,J., Furthmayr,H. and Timpl,R. (1982) *Eur. J. Biochem.*, **123**, 63-72.
Privalov,P.L. (1982) *Adv. Protein Chem.*, **35**, 1-104.
Rao,C.N., Margulies,I.M.K., Tralka,T.S., Terranova,V.P., Madri,J.A. and Liotta,L.A. (1983) *J. Biol. Chem.*, **257**, 9740-9744.
Roberts,R.D.B. and Geisow,M. (1980) *Int. J. Biol. Macromol.*, **2**, 387-389.
Rohde,H., Bachinger,H.P. and Timpl,R. (1980) *Hoppe-Seyler's Z. Physiol. Chem.*, **361**, 1651-1660.
Sakashita,S., Engvall,E. and Ruoslahti,E. (1980) *FEBS Lett.*, **116**, 243-246.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.
Smith,D.R. and Calvo,J.M. (1980) *Nucleic Acids Res.*, **10**, 2255-2274.
Sodek,J., Hodges,R.S., Smillie,L.B. and Jurasek,L. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 3800-3804.
Solter,D., Shevinsky,L., Knowles,B.B. and Strickland,S. (1979) *Dev. Biol.*, **70**, 515-521.
Strickland,S., Smith,K.K. and Marotti,K.R. (1980) *Cell*, **21**, 347-355.
Terranova,V.P., Rohrbach,D.H. and Martin,G.R. (1980) *Cell*, **22**, 719-726.
Timpl,R., Rohde,H., Robey,P.G., Rennard,S.I., Foidart,J.-M. and Martin, G.R. (1979) *J. Biol. Chem.*, **254**, 9933-9937.
Timpl,R., Johansson,S., van Delden,V., Oberbaumer,I. and Hook,M. (1983) *J. Biol. Chem.*, **258**, 8922-8927.
Vlodavsky,I. and Gospodarowicz,D. (1981) *Nature*, **289**, 304-306.
Wan,Y.-J., Wu,T.-C., Chung,A.E. and Damjanov,I. (1984) *J. Cell Biol.*, **98**, 971-979.
Wang,S.-Y. and Gudas,L.J. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 5880-5884.
Weaver,R.F. and Weissman,C. (1979) *Nucleic Acids Res.*, **7**, 1175-1193.
Wu,T.-C., Wan,Y.-J., Chung,A.E. and Damjanov,I. (1983) *Dev. Biol.*, **100**, 496-505.
Yamada,K.M. (1983) *Annu. Rev. Biochem.*, **52**, 761-799.

Received on 26 June 1984