The EMBO Journal vol.5 no.7 pp.1465–1472, 1986

Evidence from molecular cloning that SPARC, a major product of mouse embryo parietal endoderm, is related to an endothelial cell 'culture shock' glycoprotein of M_r 43 000

Ivor J.Mason, Amanda Taylor, Jeffrey G.Williams[1], Helene Sage[2] and Brigid L.M.Hogan

Laboratory of Molecular Embryology, National Institute for Medical Research, Mill Hill, London NW7 1AA, and [1]Imperial Cancer Research Fund, Clare Hall Laboratories, South Mimms, Potters Bar, Herts EN6 3LD, UK, and [2]Department of Biological Structure, University of Washington, Seattle, Washington 98195, USA

Communicated by D.A.Rees

We describe the molecular cloning and characterization of a secreted, acidic, cysteine-rich glycoprotein (SPARC) of apparent M_r 43 000 which is a major product of mouse embryo parietal endoderm. These cells are specialized for the synthesis of a rapidly expanding basement membrane, but SPARC is not itself an integral matrix component. We show that SPARC is related structurally and antigenically to an M_r 43 000 glycoprotein secreted in large amounts by bovine aortic endothelial cells as part of a 'culture shock' response to *in vitro* conditions promoting their proliferation and migration.

Key words: cDNA/43K glycoprotein/SPARC/endothelial cell/parietal endoderm

Introduction

The parietal yolk sac is one of the three membranes surrounding the mouse embryo. It consists of a homogeneous population of parietal endoderm (PE) cells which synthesize a thick basement membrane known as Reichert's membrane. This activity has made PE a favourable model system for studying the expression of genes for the structural components of basement membranes, including type IV collagen, laminin, entactin and heparan sulfate proteoglycan (Hogan *et al.*, 1982; Kurkinen *et al.*, 1983a, b; Barlow *et al.*, 1984). The parietal yolk sac has the additional feature of being a dynamic tissue, which undergoes extensive remodelling throughout post-implantation development. As the embryo grows, the yolk sac increases rapidly in surface area and expands into the uterine stroma without creating blood clots or tissue debris that would impede the filtration of fluid through Reichert's membrane. During this process the matrix first thickens, then becomes thinner and eventually breaks down before the end of pregnancy (Clark *et al.*, 1975; Dickson, 1979). The expansion of the yolk sac is accompanied by the continual recruitment of cells onto Reichert's membrane by migration of individual vimentin-positive parietal endoderm cells from the margins of a vimentin-negative epithelial layer bordering the yolk sac (Lane *et al.*, 1983; Hogan and Newman, 1984). These features suggest that further studies on the gene expression of parietal endoderm will be relevant to a variety of *in vivo* processes which involve dynamic interaction between cells and the extracellular matrix, for example during tissue morphogenesis, angiogenesis or wound-related repair. In this context it has already been shown that PE, like many other migratory cells and tissues undergoing remodelling, secretes abundant amounts of the protease, tissue plasminogen activator (Strickland *et al.*, 1976, 1980; Marotti *et al.*, 1982; see Mullins and Rohrlich, 1983, for review).

Interest has also focused on the differentiation of PE from a bipotential stem cell population, the primitive endoderm, which first appears at about 4.5 days of development in the mouse. This differentiation can be mimicked *in vitro* by the addition of retinoic acid and cAMP to F9 teratocarcinoma stem cells, providing a model system for following the co-ordinated regulation of a set of genes in embryonic cells (Strickland *et al.*, 1980; Hogan *et al.*, 1983 for review). In order to identify genes regulated in this way and expressed at high levels in PE cells, differentiated F9 cell and PE cDNA libraries have been differentially screened for sequences expressed at higher levels than in undifferentiated F9 cells. Some of the cDNAs selected in this way provided the first clones for structural components of basement membranes, including type IV collagen and laminin B chains (Kurkinen *et al.*, 1983a, b; Wang and Gudas, 1983; Barlow *et al.*, 1984; Wang *et al.*, 1985). We now report the characterization of a novel and abundant PE product — a secreted glycoprotein of M_r 43 000 (SPARC) — which does not appear to be a structural component of the extracellular matrix. We show here that SPARC is related to a glycoprotein (43K) secreted in increased amounts by bovine aortic endothelial cells as part of a co-ordinated 'culture shock' response to *in vitro* conditions stimulating their proliferation and migration (Sage, 1986; Sage *et al.*, 1984, 1986). This finding supports the idea that SPARC expression may be associated with a variety of morphogenetic processes which involve cell migration, and synthesis and modelling of the extracellular matrix.

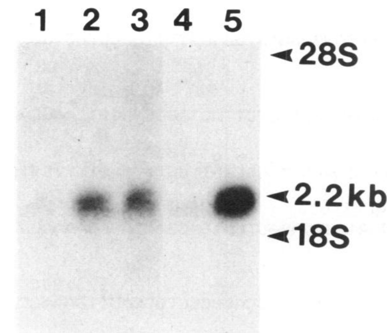

Fig. 1. Differential expression of a 2.2 kb RNA in extra-embryonic endoderm and F9 teratocarcinoma cells. Clone pF9.33 (nt 1858–2070; see Figure 2) was hybridized to Northern blots of 10 μg of total RNA from F9 undifferentiated embryonal carcinoma cells **(lane 1)**; F9 cells treated with retinoic acid and dibutyryl cAMP for 5.5 days **(lane 2)**; PYS cells **(lane 3)**. In a separate experiment pF9.33 was hybridized to 10 μg of total RNA from visceral yolk sac **(lane 4)** and parietal endoderm **(lane 5)**. The size of the hybridizing transcript was estimated from its position relative to the 18S and 28S ribosomal RNA (Wang and Gudas, 1983). From densitometer scans of lanes 1 and 2, it was calculated the level of 2.2 kb RNA increases about 20-fold after treatment of F9 cells with retinoic acid and cAMP.

© IRL Press Limited, Oxford, England

50 100
GCATTCCTGCAGCCCTTCAGACCGCCAGAACTCTTCTGCCGCCTGCCTGCCTGCCTGCCTGCCTGCCTGTGCCGAGAGTTCCCAGCATC ATG AGG GCC TGG ATC TTC TTT CTC
Met Arg Ala Trp Ile Phe Phe Leu
1

150 200
CTT TGC CTG GCC GGG AGG GCC CTG GCA GCC CCT CAG CAG ACT GAA GTT GCT GAG GAG ATA GTG GAG GAG GAA ACC GTG GTG GAG GAG ACA
Leu Cys Leu Ala Gly Arg Ala Leu Ala Ala Pro Gln Gln Thr Glu Val Ala Glu Glu Ile Val Glu Glu Glu Thr Val Val Glu Glu Thr
10 20 30

250
GGG GTA CCT GTG GGT GCC AAC CCA GTC CAG GTG GAA ATG GGA GAA TTT GAG GAC GGT GCA GAG GAA ACG GTC GAG GAG GTG GTG GCT GAC
Gly Val Pro Val Gly Ala Asn Pro Val Gln Val Glu Met Gly Glu Phe Glu Asp Gly Ala Glu Glu Thr Val Glu Glu Val Val Ala Asp
40 50 60

300 350
AAC CCC TGC CAG AAC CAT CAT TGC AAA CAT GGC AAG GTG TGT GAG CTG GAC GAG AGC AAC ACC CCC ATG TGT GTG TGC CAG GAC CCC ACC
Asn Pro Cys Gln Asn His His Cys Lys His Gly Lys Val Cys Glu Leu Asp Glu Ser Asn Thr Pro Met Cys Val Cys Gln Asp Pro Thr
70 80 90

400 450
AGC TGC CCT GCT CCC ATT GGC GAG TTT GAG AAG GTA TGC AGC AAT GAC AAC AAG ACC TTC GAC TCT TCC TGC CAC TTC TTT GCC ACC AAG
Ser Cys Pro Ala Pro Ile Gly Glu Phe Glu Lys Val Cys Ser Asn Asp Asn Lys Thr Phe Asp Ser Ser Cys His Phe Phe Ala Thr Lys
100 110 ▲ 120

500 550
TGC ACC CTG GAG GGC ACC AAG AAG GGC CAC AAG CTC CAC CTG GAC TAC ATC GGA CCA TGC AAA TAC ATC GCC CCC TGC CTG GAT TCC GAG
Cys Thr Leu Glu Gly Thr Lys Lys Gly His Lys Leu His Leu Asp Tyr Ile Gly Pro Cys Lys Tyr Ile Ala Pro Cys Leu Asp Ser Glu
130 140 150

600 650
CTG ACC GAA TTC CCT CTG CGC ATG CGT GAC TGG CTC AAA AAT GTC CTG GTC ACC TTG TAC GAG AGA GAT GAG GGC AAC AAC CTC CTC ACT
Leu Thr Glu Phe Pro Leu Arg Met Arg Asp Trp Leu Lys Asn Val Leu Val Thr Leu Tyr Glu Arg Asp Glu Gly Asn Asn Leu Leu Thr
160 170 180

700
GAG AAG CAG AAG CTG CGT GTG AAG AAG ATC CAT GAG AAT GAG AAG CGC CTG GAG GCT GGA GAC CAC CCC GTG GAG CTG TTG GCC CGA GAC
Glu Lys Gln Lys Leu Arg Val Lys Lys Ile His Glu Asn Glu Lys Arg Leu Glu Ala Gly Asp His Pro Val Glu Leu Leu Ala Arg Asp
190 200 210

750 800
TTT GAG AAG AAC TAC AAT ATG TAC ATC TTC CCT GTC CAC TGG CAG TTT GGC CAG CTG GAT CAG CAC CCT ATT GAT GGG TAC CTG TCC CAC
Phe Glu Lys Asn Tyr Asn Met Tyr Ile Phe Pro Val His Trp Gln Phe Gly Gln Leu Asp Gln His Pro Ile Asp Gly Tyr Leu Ser His
220 230 240

850 900
ACT GAG CTG GCC CCA CTG CGT GCT CCC CTC ATC CCC ATG GAA CAT TGC ACC ACA CGT TTC TTT GAG ACC TGT GAC CTA GAC AAC GAC AAG
Thr Glu Leu Ala Pro Leu Arg Ala Pro Leu Ile Pro Met Glu His Cys Thr Thr Arg Phe Phe Glu Thr Cys Asp Leu Asp Asn Asp Lys
250 260

950 1000
TAC ATT GCC CTG GAG GAA TGG GCC GGC TGC TTT GGC ATC AAG GAG CAG GAC ATC AAC AAG GAT CTG GTG ATC TAA GTTCACGCCTCCTGCTGCA
Tyr Ile Ala Leu Glu Glu Trp Ala Gly Cys Phe Gly Ile [Lys Glu Gln Asp Ile Asn Lys Asp Leu Val Ile] *
280 290 302

1050 1100
GTCCTGAACTCTCTCCCTCTGATGTGTCACCCCTCCCATTACCCCCTTGTTTAAAATGTTTGGATGGTTGGCTGTTCCGCCTGGGGATAAGGTGCTAACATAGATTTAACTGAATACATT

1150 1200 1250
AACGGTGCTAAAAAAAAAAAAAAAACAAAGTAAGAAAGAAACTAGAACCCAAGTCACAGCATTTTCCCACATAACTCTGAGGCCATGGCCCATCCACAGCCTCCTGGTCCCCTGCACTAC

1300 1350
CCAGTGTCTCACTGGCTGTGTTGGAAACGGAGTTGCATAAGCTCACCGTCCACAAGCACGAGATATCTCTAGCTTTCATTTCAATTTTGCATTTGACTCTTAACACTCACCCAGACTCTG

1400 1450
TGCTTATTTCATTTTGGGGGATGTGGGCTTTTTCCCCTGGTGGTTTGGAGTTAGGCAGAGGGAAGTTACAGACACAGGTACAAAATTTGGGTAAAGATACTGTGAGACCTGAGGACCCAC

1501 1550 1600
CAGTCAGAACCCACATGGCAAGTCTTAGTAGCCTAGGTCAAGGAAAGACAGAATAATCCAGAGCTGTGGCACACATGACAGACTCCCAGCAGCCCGGGACCTTGCTGTCTTCTCGACTCT

1650 1700
TCGGGCGTTTCTTTCCATGTTTGGCTGTTGGTTTTAGTTTTGGTGAGCCATGGGTGGGCCAGAACATCACTCAACTGCAATTGGGCTTTCAGGTTCTTGCCGGGAGCTCTAGGCACTGGG

1750 1800 1850
AGGCTGTTTCAGGAAAGTGAGACTCAAGAGGAAGACAGAAAAGGTTGTAACGTAGAGGAAGTGAGACTGGTGAATTGGTTTGATTTTTTTCACATCTAGATGGCTGTCATAAAGTTTCTA

1900 1950
GCATGTTCCCCCTCACCTCTCCCCACCCCCTGCCACTTGAAACCTTCTACTAATCAAGAGAAACTTCCAAGCCAACGGAATGGTCAGATCTCACAGGCTGAGAAATTGTTCCCCTCCAAG

2000 2050
CATTTCATGAAAAAGCTGCTTCTCATTAACCATGCAAACTCTCACAGCGATGTGAAGAGCTTGACAAGTCTTTCAAAATAAAAAGTAACAACTTAGAAACGG

Fig. 2. Nucleotide sequence and predicted amino acid sequence of differentially expressed 2.2 kb RNA. Sequence was obtained from overlapping cDNAs from PE libraries. Extension on differentiated F9 poly(A)$^+$ RNA gave a major product of 148 nt from a primer complementary to nt 127–148 (underlined). The extension product was purified and the sequence confirmed. The sequences of cDNA clones and the primer extension product from differentiated F9 mRNA were identical to those of PE cell clones, with the exception that the poly(A) tail was added after nucleotide 2070. The 11 C-terminal amino acids used to raise the antiserum are boxed. There is a glycosylation signal at amino acid 115 (arrow).

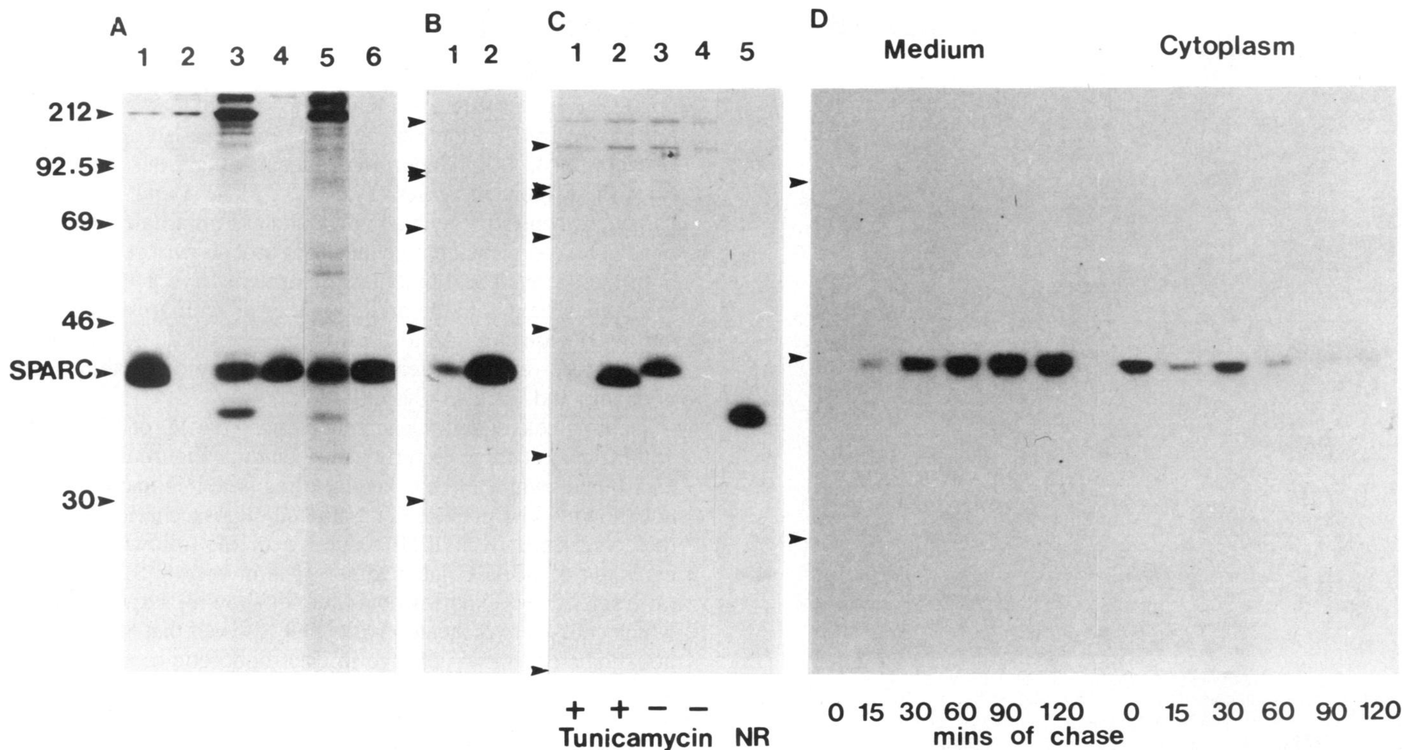

Fig. 3. Synthesis of SPARC by parietal endoderm cells and cell lines. (A) Specific immunoprecipitation of SPARC by anti-peptide serum. Approximately 4 × 10^6 parietal endoderm cells attached to 40 Reichert's membranes were labelled for 14.5 h with 50 μCi [^{35}S]methionine and 1 × 10^6 PYS cells were labelled for 16 h. **(Lane 1)** Protein immunoprecipitated from 1/8 of the culture medium of parietal endoderm using 3 μl of anti-peptide serum; **(lane 2)** as for lane 1 but antiserum pre-incubated with free peptide; **(lane 3)** 1/200 of total culture medium of parietal endoderm. From densitometer scanning of a shorter exposure of this lane it was calculated that SPARC represents about 25% of the total [^{35}S]methionine-labelled proteins accumulated in the medium during the incubation period; **(lane 4)** protein immunoprecipitated from 1/4 of the detergent lysate of parietal endoderm cells; **(lane 5)** 1/200 total culture medium of PYS cells; **(lane 6)** protein immunoprecipitated from 1/8 of the culture medium of PYS cells using 3 μl of anti-peptide serum. All samples were analysed on a 10% gel and the position of the internal ^{14}C-labelled markers is shown on the left-hand margin. **(B)** Synthesis of SPARC by F9 teratocarcinoma cells. Undifferentiated F9 cells and cells incubated for 5 days with retinoic acid and cAMP were labelled with 50 μCi [^{35}S]methionine for 10 h. Aliquots of the culture medium containing equal amounts of trichloroacetic acid-precipitable radioactivity were immunoprecipitated with 3 μl anti-peptide serum and samples analysed in a 10% gel under reducing conditions. **(Lane 1)** Undifferentiated cells; **(lane 2)** differentiated cells. From densitometer scanning of the fluorogram it was calculated that in this experiment twenty times more newly synthesized SPARC was recovered from the culture medium of differentiated compared with undifferentiated cells. All samples were analysed on a 10% gel. **(C)** Effect of tunicamycin on SPARC. F9-ACc19 cells were pre-incubated for 2 h in medium with or without 4 μg/ml tunicamycin before labelling for 4.5 h with 100 μCi [^{35}S]methionine. SPARC was immunoprecipitated from equal aliquots of the cell lysate using 5 μl of anti-peptide serum and analysed on a 10% gel under reducing conditions. **(Lanes 1 and 2)** Cells incubated with tunicamycin; **(lanes 3 and 4)** control cells; **(lanes 1 and 4)** immunoprecipitation carried out in the presence of free peptide; **(lane 5)** as for lane 3 except that cells were incubated with 50 μCi [^{35}S]methionine for 6.5 h and samples analysed under non-reducing conditions. All samples were analysed on a 10% gel. **(D)** Pulse-chase analysis of SPARC. PYS cells were pre-incubated for 45 min in medium without methionine before labelling for 15 min with 100 μCi [^{35}S]methionine. The cells were then washed and incubated in medium with excess cold methionine. SPARC was recovered by immunoprecipitation from the medium and cell cytoplasm immediately after washing (0) and after 15, 30, 60, 90 and 120 min of chase. Samples were analysed on a 7.5 − 10% gradient gel.

Results

Isolation and sequencing of overlapping cDNAs from an abundant PE transcript

Differential screening of a cDNA library from differentiated F9 cells with parietal endoderm (PE) and visceral endoderm cDNA yielded three clones (pF9.33, 52 and 54) which hybridized to a 2.2-kb polyadenylated RNA which is preferentially expressed in PE cells and parietal endoderm cell lines (Figure 1). Calculations show that this RNA represents about 0.5% of the polyadenylated RNA in PE cells and is more abundant than either of the laminin B chains (I.Mason, unpublished results). Undifferentiated F9 cells have only very low levels of the RNA, but it accumulates 20-fold, to 0.15% of the polyadenylated RNA after treatment with retinoic acid and cAMP (Figure 1, lanes 1 and 2; Mason *et al.*, 1986).

Sequence analysis revealed that these clones were all derived from the 3′ end of the same mRNA. They were then used to probe two PE cDNA libraries, and the nucleotide sequence derived from several of the longest cDNAs isolated is shown in Figure 2. Primer extension, using a single stranded primer derived from nucleotides (nt) 127 − 148, gave a product of 148 nt. This was isolated and sequenced, and confirmed the isolation of full-length cDNA clones. Conceptual translation of the sequence gave a unique open reading frame between the first AUG (nt 90) and the stop codon TAA (nt 996). This was followed by 1083 nt of 3′ non-coding sequence, a polyadenylation signal (nt 2054 − 2059) and a poly(A) tail.

Primary structure of the predicted protein

Analysis of the deduced sequence of 302 amino acids revealed at least four distinct protein domains. Domain I, residues 1 − 17, presents a classical signal sequence (von Heijne, 1985), strongly suggesting that the protein is secreted. A probable site of cleavage is predicted by a sharp change in hydrophobicity around residues 18 − 20. Domain II, residues 23 − 68, is rich in acidic amino acids, particularly glutamic acid (15/46). Analysis of this

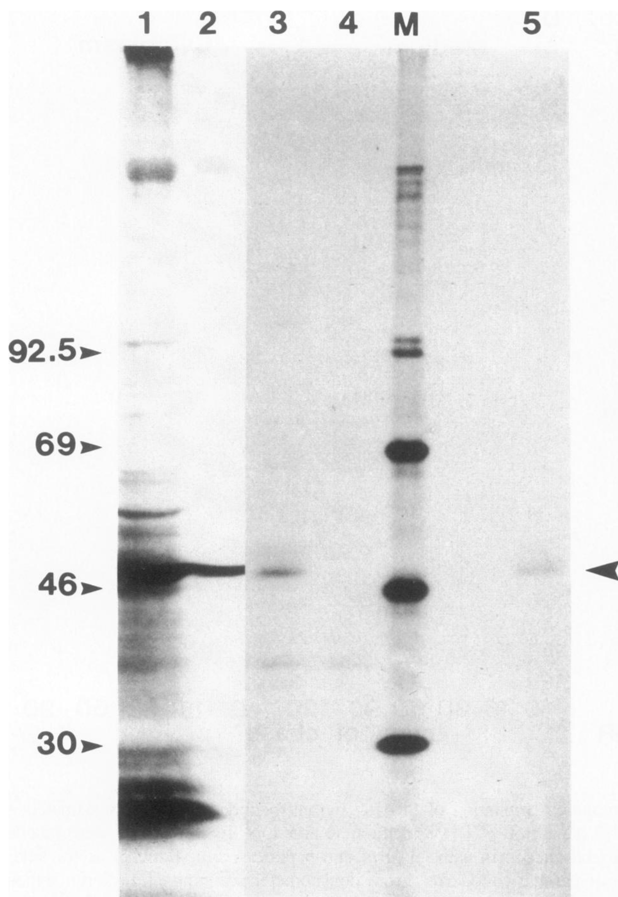

Fig. 4. Hybrid selection and *in vitro* translation of SPARC mRNA. RNA from differentiated F9 cells was translated in the nuclease-treated reticulocyte lysate and total products or proteins were immunoprecipitated with anti-peptide serum and subsequently analysed on a 6–12% gel under reducing conditions. **(Lane 1)** Total proteins translated from 0.2 μg poly(A)$^+$ RNA after removal of RNA hybridizing to pPE220; **(lane 2)** control, without added RNA; **(lane 3)** protein recovered by anti-peptide serum from lysate primed with 4 μg total RNA; **(lane 4)** as for lane 3 but antiserum pre-incubated with free peptide; **(lane 5)** 17 μg poly(A)$^+$ RNA was hybridized to pPE220. Bound RNA was eluted and *in vitro* translation product immunoprecipitated with anti-peptide serum. Of the remaining RNA, 0.2 μg was used in the translation shown in lane 1. The large arrowhead indicates the M_r = 47 500 *in vitro* translation product.

sequence by the method of Chou and Fasman (1978) gives a strong prediction for alpha helix. Domain III is a central region of ~100 residues (69–170) containing 11 cysteine residues and the unique glycosylation site at residue 115. The remaining domain, IV, encompasses residues 171–302 and contains approximately equal numbers of acidic (21) and basic (22) residues and three cysteines near the C-terminus.

These data predict that PE cells and differentiated F9 cells synthesize large amounts of a secreted acidic glycoprotein rich in S–S bonds. To test these predictions, a peptide corresponding to the C-terminal sequence Lys-Glu-Glu-Asp-Ile-Asn-Lys-Asp-Leu-Val-Ile was synthesized, coupled to keyhole limpet haemocyanin (KLH) and used to raise antibodies in rabbits.

Peptide antibodies recognise a M_r 43 000 secreted glycoprotein

The anti-peptide serum specifically recognizes one of the major ^{35}S-labelled proteins secreted by PE cells, the parietal endoderm cell line PYS and the differentiated F9 cell line F9-ACc19. This protein represents about 25% of the total [^{35}S]methionine labelled proteins accumulating in the culture medium of PE cells over 14 h (Figure 3A, lane 3). Moreover, differentiated F9 cells produce between 20- and 40-fold more of this protein than undifferentiated cells (Figure 3B, lanes 1 and 2; and unpublished results). From its migration relative to internal markers, the protein recognized by the anti-peptide serum has an apparent M_r of 40 000–43 000. Under non-reducing conditions the mobility of the protein increases (Figure 3C, lane 5 and Table I), a behaviour consistent with a high content of intramolecular S–S bonds. Two-dimensional gel analysis (not shown) revealed that both the secreted and cytoplasmic protein have a Pi of 4.3, a result consistent with the predominance of acidic residues in the predicted sequence. As a result of these properties, we subsequently refer to this protein as SPARC (Secreted Protein which is Acidic and Rich in Cysteines).

Treating cells with tunicamycin reduced the M_r of SPARC by 1300 (mean of three experiments including Figure 3C, lanes 2 and 3) and completely blocked labelling with [^{3}H]mannose (data not shown). The presence of N-linked oligosaccharide was further confirmed by a 1250 reduction in M_r following endo-H digestion of SPARC labelled for 15 min with [^{35}S]methionine (average of two experiments, data not shown). Experiments in which cells were incubated with $^{35}SO_4$ showed that SPARC was not sulphated, and no change in electrophoretic mobility of the secreted protein was observed during a pulse-chase experiment over 2 h (Figure 3D).

To confirm that the anti-peptide serum specifically recognizes the translation product encoded by the sequence shown in Figure 2, RNA from differentiated F9 cells was translated in the nuclease-treated reticulocyte lysate. The anti-peptide serum specifically immunoprecipitated a single protein of M_r 47 500 (Figure 4, lanes 3 and 4), and the same protein was recovered in hybrid-select translation experiments using the cDNA pPE 220 which spans nt 571–1160 (Figure 4, lane 5). No phosphorylation of the *in vitro* translated 47 500 protein was observed when [^{32}P]ATP was included in the translation reaction, although other proteins were labelled (data not shown).

SPARC is not an integral basement membrane component

Indirect immunofluorescence with anti-peptide serum failed to stain either the extracellular matrix of PYS cells or Reichert's membrane of 8.5 day p.c. mouse embryos under conditions where a strong positive reaction was obtained with anti-laminin/entactin serum (data not shown). However, this result might be due to masking of the C-terminus of the protein if it is incorporated into the matrix. The following experiment therefore tested the possibility that SPARC is incorporated into the basement membrane assembled by PE cells. Reichert's membranes were dissected from 12.5 day p.c. embryos and incubated with [^{35}S]methionine for 16 h. The membranes were then washed and incubated for 90 min in an excess of cold methionine to chase radioactive SPARC from the cytoplasm before lysing the cells with detergent. Aliquots of both total membrane and culture medium were solubilized in SDS and β-mercaptoethanol at 100°C, and the proteins separated by SDS gel electrophoresis and transferred to nitrocellulose. The blots were then probed with anti-peptide serum followed by autoradiography. As shown in Figure 5A, lanes 1 and 2, the ratio of radioactive SPARC in the medium to membrane was ~60:1. Moreover, by the end of the 16 h incubation more SPARC had accumulated in the medium than was incorporated into the membranes during the whole of the preceding period of development *in vivo* (Figure 5A, lanes 3 and 4). Control experiments showed a very different distribution pattern for laminin, with approximately four times more radioactive laminin in the membranes than in the medium (Figure

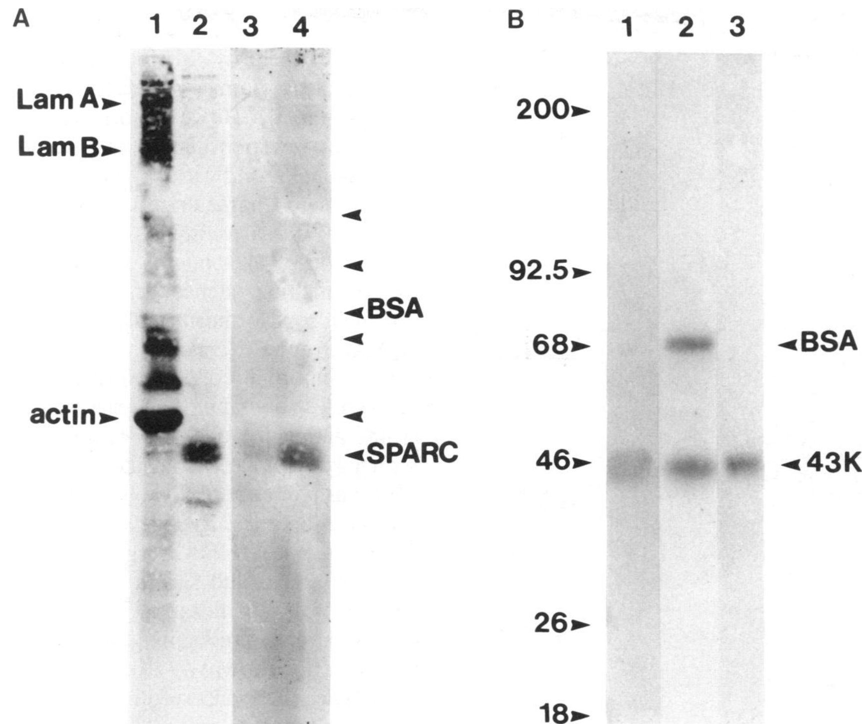

Fig. 5 Western blotting with anti-peptide serum. **(A)** Relative distribution of SPARC between medium and Reichert's membrane fractions. Fourteen Reichert's membranes with attached PE cells from 12.5-day embryos were incubated for 16 h in 1 ml medium containing 1 μg/ml cold methionine and 50 μCi [^{35}S]methionine. The membranes were washed and incubated for a further 90 min in medium with excess cold methionine, followed by further washes, lysis of the cells, and solubilization of the matrix and cell ghosts in twice-concentrated gel sample buffer. Under the conditions of labelling and chase, any radioactive protein recovered from the basement membrane fraction is likely to have been incorporated into the structure of the matrix rather than non-specifically bound after cell lysis. Fractions of 1/80 of the total membrane fraction and 1/300 of the total medium were analysed on a 10% gel under reducing conditions. **(Lane 1)** Autoradiogram of Reichert's membrane blot shown in lane 3. The major band of M_r 46 000 is actin from the cell ghosts. The position of the laminin A and B chains is indicated by arrows; **(lane 2)** autoradiogram of the blot of medium proteins shown in lane 4. From densitometer scans of the SPARC band in lanes 1 and 2 and the relative amount of each fraction analysed on the gel, the ratio of radioactive SPARC in the medium to membrane was calculated to be 60:1, **(lane 3)** SPARC in the membrane fraction revealed by Western blotting using anti-peptide serum and peroxidase-labelled second antibody; **(lane 4)** SPARC in the medium fraction revealed by Western blotting. The position of major unstained proteins, including bovine serum albumin (BSA), is indicated by arrows in the right-hand margin. Control experiments showed that no SPARC was present in the medium before incubation. **(B)** Immunological cross-reaction between anti-peptide serum and bovine 43 000 protein. Partially purified [^{3}H]proline-labelled 43 000 protein from the culture medium of BAE cells was separated on a 10% gel and transferred to nitrocellulose. **(Lane 1)** reaction with anti-peptide serum and peroxidase-labelled second antibody; **(lane 2)** transferred proteins stained with amido black; **(lane 3)** autoradiogram of lane 2 showing that only the 43 000 protein was synthesized by the BAE cells. The position of internal stained markers is shown in the left-hand margin and the position of the 43 000 protein and co-purifying bovine serum albumin in the right-hand margin.

5A, lanes 1 and 2, and data not shown). We conclude that, unlike laminin, SPARC is not an integral basement membrane component.

Parietal endoderm SPARC is related structurally and antigenically to bovine endothelial cell 43 000 glycoprotein

No significant homology was found between SPARC and any protein or nucleotide sequence in the NIH or EMBL data bases. However, it has many properties in common with an albumin-binding glycoprotein (43K protein) originally recovered from the conditioned medium of bovine aortic endothelial (BAE) cells where it constitutes ~1–3% of total biosynthetically labelled medium protein (Sage *et al.*, 1984). In a survey of cell lines, these authors found that the teratocarcinoma-derived parietal endoderm cell line PF-HR9 secreted a protein with similar chromatographic properties and peptide maps to the endothelial cell 43K protein. The antiserum raised against the synthetic peptide of mouse SPARC was therefore tested for its ability to recognize specifically 43K made by bovine aortic endothelial (BAE) cells and foetal bovine ligamentum nuchae fibroblasts (FCLF; Sage, 1985). In both cases positive results were obtained (Figure 5B and data not shown).

One of the characteristic features of endothelial cell 43K protein is its limited sensitivity to prolonged trypsin digestion, with the production of a major glycosylated fragment of M_r ~30 000 under reducing conditions (Sage *et al.*, 1984). A very similar pattern of trypsin cleavage was seen with [^{35}S]methionine- and [^{3}H]mannose-labelled SPARC from PYS cells (Figure 6).

Finally, a unique N-terminal sequence Ala-Pro-Gln-Gln-Glu-Ala-Leu-Pro-Asp-Glu-Cys-X-Val-Val was obtained directly from purified bovine 43K protein. The first four residues of this sequence are identical to residues 18–21 immediately following the predicted signal sequence of the murine protein. Figure 7 shows the best alignment of the remaining bovine and predicted murine sequences.

Attempts to show albumin-binding activity in [^{35}S]methionine-labelled SPARC by treating PYS culture medium with BSA-Sepharose, have been unsuccessful. However, all the SPARC secreted by the PYS cells could already be complexed with albumin or a unique M_r 70 000 protein derived from the serum present in the culture medium (Sage, 1986).

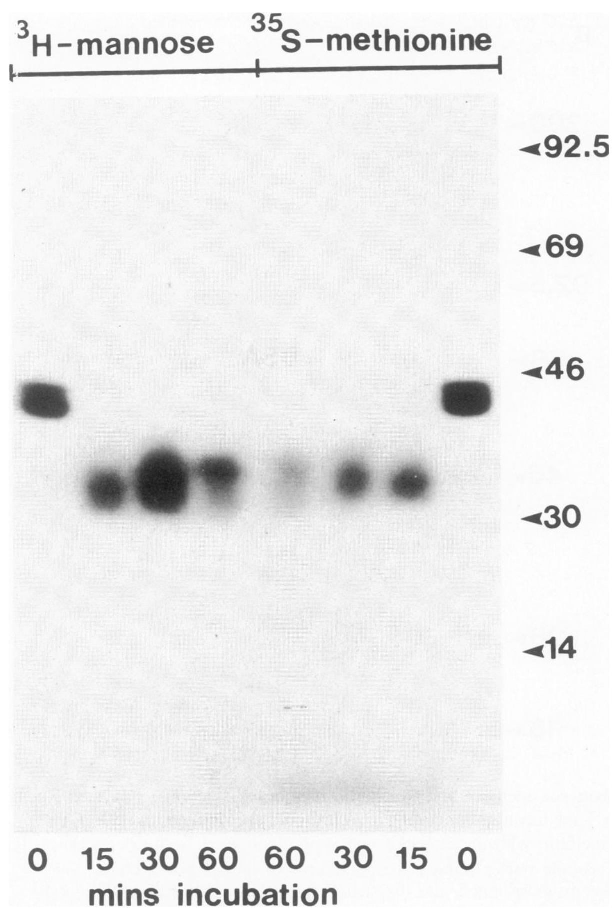

Fig. 6. Trypsin cleavage of SPARC. PYS cells were incubated for 4 h with 100 μCi [^{35}S]methionine or 20 h with 1 mCi/ml [^{3}H]mannose, and SPARC was immunoprecipitated from aliquots of the culture medium with 3 μl anti-peptide serum and protein A Sepharose. Immunocomplexes were incubated with trypsin for 0, 15, 30 or 60 min and then analysed on a 6–12% gradient gel under reducing conditions. Fluorography of the [^{3}H]mannose-labelled samples was for 1 month and the [^{35}S]methionine-labelled samples for 2 days.

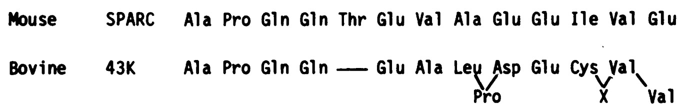

Fig. 7. Bovine 43K and predicted murine SPARC N-terminal sequences aligned for maximum homology.

Discussion

By cDNA cloning, sequence analysis, *in vitro* translation and preparation of an anti-peptide serum we have characterized a novel glycoprotein product of mouse embryo PE cells, SPARC. This protein represents ~25% of the total [^{35}S]methionine-labelled products accumulating in the culture medium over a 14.5-h labelling period (Figure 3A, lane 3). Most of the properties of SPARC are those predicted from the deduced amino acid sequence. It is a secreted, acidic glycoprotein, and behaves on gel electrophoresis as expected for a protein with many intramolecular disulphide bonds. The fact that the major tryptic fragment is labelled with [^{3}H]mannose suggests that these S–S bonds stabilize a protease-resistant core which includes the asparagine at position 115. The high content of acidic residues, especially Glu, predicted for the N-terminal domain of the protein probably accounts for the difference between the M_r calculated for the predicted protein (34 411) and the M_r of 47 500 calculated from the mobility of the *in vitro* translation product on SDS–polyacrylamide gel electrophoresis relative to marker proteins (Figure 4). However, although labelling experiments have ruled out phosphorylation and sulphation as additional post-translational modifications, we have not yet excluded all formal possibilities. Characterization of cosmid clones and analysis of genomic DNA are consistent with only a single gene for SPARC in the mouse (Mason *et al.*, 1986; J.McVey *et al.*, unpublished results) and no evidence for alternate processing of this gene was obtained by Northern analysis or from sequences of twelve cDNAs from PE cells.

One important conclusion from this work is that the SPARC produced so abundantly by parietal endoderm cells is structurally and antigenically closely related to a major 43K glycoprotein secreted by bovine endothelial cells and foetal ligamentum nuchae fibroblasts in culture. The two proteins have similar chromatographic properties, tryptic cleavage patterns and amino acid compositions, and are not integral components of the extracellular matrix (Table I; and Sage *et al.*, 1984, 1986; Sage, 1985, 1986). Moreover, antibodies to a C-terminal peptide of the mouse protein specifically recognize the 43K bovine protein (Figure 5B and data not shown) — a result which could only be obtained if sequences are conserved in this region. Finally, we have shown that the first four amino acid residues at the N-terminus of the bovine protein are identical to four residues at the predicted cleavage site of the signal sequence of the mouse protein, and that other identical or related residues follow.

Although no specific function has yet been ascribed to the endothelial 43K protein, previous work has ruled out identity to a wide range of serum proteins and to a variety of known proteases and protease inhibitors, including serine proteases, collagenase (including latent collagenase), and plasminogen activator and its inhibitors (Sage *et al.*, 1984; 1986). The high level of 43K in the culture medium probably excludes a function as a growth factor, and it is not assembled as a structural component into the extracellular matrix of bovine endothelial cells or fibroblasts in culture (Sage *et al.*, 1984, 1986). However, 43K does have the property of forming a complex with two serum proteins, bovine serum albumin (Sage *et al.*, 1984) or, probably more specifically, a novel serum protein of M_r 70 000 (Sage, 1986). It appears that this 70 000 protein is selectively taken up by endothelial cells from the culture medium, and is rapidly released in a complex with 43K when the cells are exposed to serum-free conditions (Sage, 1986). Most significantly, the rate of synthesis of 43K protein is very dependent on culture conditions. For example, when confluent BAE cells are seeded at low density in serum-containing medium there is a 10-fold increase in the specific synthesis of 43K protein and a 3-fold increase when confluent cultures in the absence of serum are exposed to doses ($10^{-6}-10^{-10}$ g/ml) of endotoxin, which causes cells to round up and detach, but not lyse (Sage *et al.*, 1986). These and other observations have led to the idea of a 'culture shock' response, which is elicited, for example when endothelial cells are deprived of contact with their basement membrane and are plated out on a plastic surface in the presence of high concentrations of serum-derived growth factors (Sage *et al.*, 1986; Sage, 1986). The end result of the response is that cells migrate to fill the available space and resynthesize a continuous layer of extracellular matrix material beneath their basal surface. Under the conditions in which it is most usually elicited — plating out at low density in serum-containing medium — the endothelial cell

culture shock response is associated with proliferation and the expression of proliferation-related genes (Barrett *et al.*, 1984). However, increased 43K secretion is also seen in the absence of serum-derived growth factors when an increase in cell number does not occur (Sage, 1986).

We have shown here that parietal endoderm cells of the mouse embryo secrete very high levels of SPARC. They also express elevated levels of tissue plasminogen activator (Strickland *et al.*, 1976; Marotti *et al.*, 1982) and the proto-oncogene c-*fos* (I.Mason *et al.*, 1985), as well as basement membrane components such as laminin and type IV collagen. They are actively migrating from the edge of an epithelial sheet and are involved in synthesizing and remodelling the extracellular matrix. This raises the possibility that genes constitutively expressed at high levels by PE cells — including tissue plasminogen activator, c-fos, SPARC and basement membrane components — may be specifically activated when adult cells are detached from their normal substratum (the basement membrane) and are required to migrate and restore their extracellular matrix, with or without proliferation. We therefore predict that high SPARC expression may be a component of a variety of wound-related reactions as well as normal morphogenetic processes which involve cell migration on, or remodelling of, basement membranes. In this regard it is of interest that the axonal regeneration of a variety of nerves *in vivo* is associated with the induction of large amounts of acidic proteins of M_r 43 000–49 000 (Skene, 1984). Experiments designed to test these predictions and to elucidate the precise function of the protein are underway.

Materials and methods

Dissection and cell culture

Parietal and visceral yolk sacs were dissected from C3H/He mice at 13.5 days *post coitum.* Residual trophoblast cells were removed from Reichert's membrane, leaving a homogeneous population of PE cells. F9 teratocarcinoma stem cells: differentiation of F9 teratocarcinoma stem cells to parietal endoderm was achieved by adding 5×10^{-8} M all-*trans* retinoic acid (Sigma), 10^{-4} M dibutryl cAMP (Sigma) and 10^{-4} M isobutylmethylxanthine (Sigma) for 5.0–5.5 days (Cooper *et al.*, 1981). The teratocarcinoma-derived parietal endoderm cell line, PYS-2, was originally obtained from Dr J.M.Lehman (Lehman *et al.*, 1974) and F9-ACC19 from Dr D.Solter (Solter *et al.*, 1979).

RNA isolation and Northern analysis

Total cellular RNA and poly(A)$^+$ RNA were isolated as described by Kurkinen *et al.* (1983b) and Maniatis *et al.* (1982), and quantitated by spectrophotometry and [^{3}H]poly(U) assay (Bishop *et al.*, 1974). For Northern analysis, RNA was electrophoresed in 1% formaldehyde gels (Maniatis *et al.*, 1982). Gels were washed in distilled water for 5 min and the RNA transferred overnight to Gene-Screen Plus membranes (New England Nuclear) in 20 × SSC. Hybridization and washing were carried out as described by Mason *et al.* (1985).

cDNA libraries and sequencing

cDNA libraries were made from differentiated F9 and mouse embryo PE cells (Kurkinen *et al.*, 1983a, c). The differentiated F9 library was screened by differential hybridization (Williams and Lloyd, 1978) using cDNA probes from PE and visceral yolk sac RNA. Three clones (nt 1824–2070) were isolated which hybridized to a 2.2-kb RNA and were used to isolate longer overlapping clones from the PE cDNA library; the longest, pPE.30, spanned the mRNA from the poly(A) tail to an *Eco*RI site (nt 679). A second PE library was then constructed from 2.5 μg of poly(A)$^+$ RNA using, in addition to dT_{14}, 100 ng of a single stranded primer (P.J.Mason *et al.*, 1985) complementary to nt 828–1011 from a *Kpn*I–*Pst*I fragment of pPE.30. cDNA was prepared and cloned into pUC8 essentially as described by Gubler and Hoffman (1983). To isolate longer clones, the library was screened with a fragment (nt 571–828) 5' to the primer. Sequencing was carried out in both strands according to Maxam and Gilbert (1980) and Sanger *et al.* (1977) using the pEMBL and M13 vectors (Dente *et al.*, 1983; Messing *et al.*, 1981).

Primer extension

Primer extension was performed using a strand-separated primer (nt 127–148) (Williams and Mason, 1985). Sixteen nanogrammes of primer were annealed to 40 μg differentiated F9 poly(A)$^+$ RNA at 70°C for 2 h and extension carried out at 42°C for 1 h. The extension products were separated on an 8% sequencing gel and sequenced (Maxam and Gilbert, 1980). Accurate quantitation of the abundance of the 2.2-kb RNA was obtained by measuring the incorporation of known amounts of primer into S1-resistant hybrids with a known amount of poly(U)-assayed RNA (Bishop *et al.*, 1974).

Hybrid selection and in vitro translation

RNA was selected from 17 μg of poly(A)$^+$ RNA from differentiated F9 cells using a cDNA spanning nt 571–1162 (Maniatis *et al.*, 1982). The reticulocyte lysate was a gift from Dr R.J.Jackson and translation was performed as detailed by Pelham and Jackson (1976). To assay for phosphorylation of SPARC, 500 μCi [^{32}P]ATP (sp. act. >5000 Ci/mmol) was added to the reaction. A number of proteins, but not SPARC, were labelled under these conditions.

Antisera

The peptide Tyr-Ile-Lys-Glu-Gln-Asp-Ile-Asn-Lys-Asp-Leu-Val-Ile containing the 11 C-terminal amino acids predicted from the cDNA sequence and N-terminal Tyr was synthesized by the Merrifield method. The peptide was coupled to keyhole limpet haemocyanin and injected into a rabbit to raise antibodies (Harvey *et al.*, 1984; I.Mason *et al.*, 1985). The antiserum produced was active in ELISA assay against substrate-bound peptide at dilutions of 1:4000, and in immunoprecipitation assays at a dilution of 1:1000. Affinity purification was carried out by cycling a crude immunoglobulin fraction over columns of KLH–Sepharose and peptide–Sepharose. Antibodies binding to the peptide–Sepharose were eluted with 1 M propionic acid, neutralized, dialysed against 0.15 M NaCl, 0.05 M Tris, pH 7.6, and concentrated by Amicon filtration.

Labelling cells and tissues

Labelling of cells with [^{35}S]methinoine, [^{35}S]sulphate and [^{3}H]mannose, and tunicamycin and endo-H treatment have been described elsewhere (Cooper *et al.*, 1981; Hogan *et al.*, 1982; Kurkinen *et al.*, 1984). Previous experience had shown that >2.0 μg/ml tunicamycin are required to completely inhibit glycosylation of M_r 47 000 colligin in parietal endoderm cells and cell lines (Kurkinen *et al.*, 1984). In the experiments described here 2.5, 3.0 and 4.0 μg/ml tunicamycin were used with similar results, and controls with colligin included.

Purification and sequencing of 43 000 protein from bovine aortic endothelial (BAE) cells and foetal calf ligamentum nuchae fibroblasts (FCLF)

BAE cells (a gift from Dr S.Schwartz, University of Washington, Seattle, WA) and FCLF (a gift from Dr R.Mecham, Washington University, St Louis, MO) were cultured and metabolically labelled with [^{3}H]proline (Sage *et al.*, 1984; Sage, 1985). Labelled proteins were purified from the culture media by precipitation with ammonium sulphate (20–50%, w/v ratio), followed by chromatography on DEAE–cellulose (Sage *et al.*, 1984). Final purification was achieved by successive chromatographic elution from Sephadex G-200 in a 0.15 M NaCl, 0.05 M Tris–HCl buffer, pH 7.5, at room temperature. M_r 43 000 proteins from both types of cells were identical by one- and two-dimensional peptide mapping (Sage *et al.*, 1984) and were homogeneous by SDS–polyacrylamide gel electrophoresis (SDS-PAGE).

Approximately 400 pmol of bovine 43 K protein were sequenced directly in an Applied Biosystems 470 gas phase machine. PTH-amino acids were identified by reverse-phase h.p.l.c. on Waters Novapak C_{18} columns by isopropanol gradient elution. Sequencing was continued until the background due to random acid-catalysed cleavage of peptide bonds prevented unambiguous amino acid assignments.

Immunoprecipitation, gel electrophoresis, trypsin cleavage and Western blotting

Immunoprecipitation from culture medium and detergent lysates of cells was performed with 3–5 μl of anti-peptide serum and protein A coupled to either Sepharose or agarose (Cooper *et al.*, 1981). In control experiments the reaction was blocked by pre-incubation of the antiserum with an equal volume of free peptide (3 mg/ml) for 15 min at room temperature (I.Mason *et al.*, 1985). Samples were analysed by SDS-PAGE and fluorography as described (Cooper *et al.*, 1981).

Internal ^{14}C-labelled markers were myosin (M_r = 212 000), phosphorylase b (100 000 and 92 500), bovine serum albumin (69 000), ovalbumin (46 000), carbonic anhydrase (30 000) and lysozyme (14 000) (Amersham International). In some experiments SPARC migrated on SDS gels as a doublet. However, only a single band was present after reduction and alkylation, suggesting that the doublet was an artefact caused by incomplete reduction and disulphide exchange. Identical results were obtained in immunoprecipitation reactions with whole serum, or immunoglobulin purified by affinity chromatography on peptide–Sepharose.

Proteins separated by SDS-PAGE were transferred to nitrocellulose by the method of Towbin *et al.* (1979), except that the transfer buffer was 0.192 M glycine, 0.025 M Tris, pH 8.3, 20% methanol, and unbound sites were blocked with 1% haemaglobin (Hb) in PBS. Blots were incubated overnight in anti-peptide serum diluted 1:500 in 1% Hb in PBS, followed by horseradish peroxidase-linked goat anti-rabbit at 1:100 dilution, and developed in 0.05% diaminobenzidine, 0.01% H_2O_2 in PBS. Trypsin cleavage of immunoprecipitated protein bound to protein A was carried out under the conditions described by Sage *et al.* (1984).

Each sample was incubated with 1 μg trypsin (bovine pancreas Type III, Sigma) in 30 μl 0.15 M NaCl, 0.05 M Tris-HCl, pH 7.5, at 37°C and the reaction stopped by adding twice-concentrated gel sample buffer and heating to 100°C for 3 min.

Acknowledgements

We thank Denis Banville, Denise Barlow, Robb Krumlauf, John McVey, Philip Mason, David Murphy and Hugh Pelham for advice and discussion. Particular thanks are due to Markku Kurkinen for isolating the first clones and subsequent help and encouragement. The N-terminal sequence of bovine 43K was obtained by Michael Geisow, National Institute for Medical Research, and we acknowledge his many useful suggestions. Skilled technical assistance was provided by Jennifer Hanson, Lynne Eaton, Sarah Connoly and Margie Chow. The earlier stages of this work were supported by the Imperial Cancer Research Fund. H.S. is an Established Investigator of the American Heart Association and part of this work was funded by the National Institutes of Health grants GM29853 and HL18645, and by a grant from R.J.Reynolds, Inc.

References

Barlow,D.P., Green,N.M., Kurkinen,M. and Hogan,B.L.M. (1984) *EMBO J.*, **3**, 2355–2362.

Barrett,T.B., Gajdusek,C.M., Schwartz,S.M., McDougall,J.K. and Benditt,E.P. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 6772–6774.

Bishop,J.O., Rosbach,M. and Evans,D. (1974) *J. Mol. Biol.*, **85**, 75–86.

Chou,P.Y. and Fasman,G.D. (1978) *Annu. Rev. Biochem.*, **47**, 251–276.

Cooper,A.R., Kurkinen,M., Taylor,A. and Hogan,B.L.M. (1981) *Eur. J. Biochem.*, **119**, 189–197.

Clark,C.C., Minor,R.R., Koszalka,T.R., Brent,R.L. and Kefalides,N.A. (1975) *Dev. Biol.*, **46**, 243–261.

Dente,L., Cesareni,G. and Cortese,R. (1983) *Nucleic Acids Res.*, **11**, 1645–1655.

Dickson,A.D. (1979) *J. Anat.*, **129**, 571–577.

Gubler,U. and Hoffman,B.J. (1983) *Gene*, **25**, 263–269.

Harvey,R., Oostra,B.A., Belsham,G.J., Gillett,P. and Smith,A.E. (1984) *Mol. Cell. Biol.*, **4**, 1334–1342.

Hogan,B.L.M. and Newman,R. (1984) *Differentiation*, **26**, 138–143.

Hogan,B.L.M., Barlow,D.P. and Tilly,R. (1983) *Cancer Surveys*, **2**, 115–140.

Hogan,B.L.M., Taylor,A. and Cooper,A.R. (1982) *Dev. Biol.*, **90**, 210–214.

Kurkinen,M., Barlow,D.P., Helfman,D., Williams,J.G. and Hogan,B.L.M. (1983a) *Nucleic Acids Res.*, **11**, 6199–6209.

Kurkinen,M., Barlow,D.P., Jenkins,J.R. and Hogan,B.L.M. (1983b) *J. Biol. Chem.*, **258**, 6543–6548.

Kurkinen,M., Cooper,A.R., Barlow,D.P., Jenkins,J.R. and Hogan,B.L.M. (1983c) *Cold Spring Harbor Conferences on Cell Proliferation. 10. Teratocarcinoma Stem Cells.* Cold Spring Harbor Publications, Cold Spring Harbor, NY, pp. 389–401.

Kurkinen,M., Hogan,B.L.M., Taylor,A. and Garrels,J.I. (1984) *J. Biol. Chem.*, **259**, 5915–5922.

Lane,E.B., Hogan,B.L.M., Kurkinen,M. and Garrels,J.I. (1983) *Nature*, **303**, 701–703.

Lehman,J.M., Speers,W.C., Swartzendruber,D.E. and Pierce,G.B. (1974) *J. Cell. Physiol.*, **84**, 13–28.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning.* Cold Spring Harbor Publications, Cold Spring Harbor, NY.

Marotti,K.R., Belin,D. and Strickland,S. (1982) *Dev. Biol.*, **90**, 154–159.

Mason,I., Murphy,D. and Hogan,B.L.M. (1985) *Differentiation*, **30**, 76–81.

Mason,P.J., Jones,M.B., Elkington,J.A. and Williams,J.G. (1985) *EMBO J.*, **4**, 205–211.

Mason,I.J., Murphy,D., Münke,M., Francke,U., Elliott,R.W. and Hogan,B.L.M. (1986) *EMBO J.*, **5**, in press.

Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499–560.

Messing,J., Crea,J. and Seeburg,P.H. (1981) *Nucleic Acids Res.*, **9**, 309–321.

Mullins,D.E. and Rohrlich,S.T. (1983) *Biochim. Biophys. Acta*, **695**, 177–214.

Pelham,H.R.B. and Jackson,R.J. (1976) *Eur. J. Biochem.*, **67**, 247–256.

Sage,H. (1985) *Biochemistry*, **24**, 7430–7440.

Sage,H. (1986) *J. Biol. Chem.*, in press.

Sage,H., Johnson,C. and Bornstein,P. (1984) *J. Biol. Chem.*, **259**, 3993–4007.

Sage,H., Tupper,J. and Bramson,R. (1986) *J. Cell Physiol.*, in press.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.

Skene,J.H.P. (1984) *Cell*, **37**, 697–700.

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.

Strickland,S., Reich,E. and Sherman,M.I. (1976) *Cell*, **9**, 231–240.

Towbin,H., Staehelin,T. and Gordon,J. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 4350–4354.

Von Heijne,G. (1985) *J. Mol. Biol.*, **184**, 99–105.

Wang,S.-Y. and Gudas,L.J. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 5880–5884.

Wang,S.-Y., LaRosa,G.J. and Gudas,L.J. (1985) *Dev. Biol.*, **107**, 75–86.

Williams,J.G. and Lloyd,M.M. (1978) *EMBO J.*, **4**, 205–211.

Williams,J.G. and Mason,P.J. (1985) *Hybridisation in the analysis of RNA.* In Hames,B.D. and Higgins,S.J. (eds), *Nucleic Acid Hybridisation.* IRL Press, Oxford, UK, pp. 139–160.

Received on 1 April 1986