

# Close similarity of epidermal growth factor receptor and *v-erb-B* oncogene protein sequences

J. Downward\*, Y. Yarden†, E. Mayes\*, G. Scrase\*, N. Totty\*, P. Stockwell\*, A. Ullrich‡, J. Schlessinger† & M. D. Waterfield\*

\* Protein Chemistry Laboratory, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK

† Department of Chemical Immunology, The Weizmann Institute of Science, Rehovot, Israel

‡ Genentech Incorporated, 460 Point San Bruno Boulevard, San Francisco, California 94080, USA

*Each of six peptides derived from the human epidermal growth factor (EGF) receptor very closely matches a part of the deduced sequence of the v-erb-B transforming protein of avian erythroblastosis virus (AEV). In all, the peptides contain 83 amino acid residues, 74 of which are shared with v-erb-B. The AEV progenitor may have acquired the cellular gene sequences of a truncated EGF receptor (or closely related protein) lacking the external EGF-binding domain but retaining the transmembrane domain and a domain involved in stimulating cell proliferation. Transformation of cells by AEV may result, in part, from the inappropriate acquisition of a truncated EGF receptor from the c-erb-B gene.*

REGULATION of the proliferation of cells in culture can be influenced by a number of mitogens including a series of polypeptide growth factors which, acting alone or synergistically with other mitogens, can induce DNA synthesis and proliferation of specific target cells (for recent reviews see ref. 1). Epidermal growth factor (EGF) and platelet-derived growth factor (PDGF) are probably the best characterized growth factors: the precise function of these polypeptides *in vivo* is, however, unclear. EGF may have a role in cell proliferation and differentiation since it will induce early eyelid opening and incisor development in new born mice<sup>2</sup>; PDGF on the other hand, which is released from platelets during blood clot formation at wound sites, may have a role in repair processes<sup>3</sup>. These and other growth factors *in vitro* can trigger a variety of morphological and biochemical changes that resemble those characteristic of transformed cells, and have also been implicated in the abnormal regulation of proliferation shown by transformed and tumour-derived cell lines (reviewed in refs 4 and 5). Thus it has been suggested that transformed cells may both synthesise and respond to growth factors and consequently proliferate independently through 'autocrine' secretion<sup>6</sup>. Direct support for such an autocrine role for aberrantly expressed growth factors in the control of abnormal cell proliferation came recently from the discovery that the putative transforming protein (p28<sup>sis</sup>) of simian sarcoma virus (SSV) is structurally related to the growth factor PDGF<sup>7-9</sup> and can also function like PDGF as a growth factor for cells in culture<sup>10</sup>. Other growth factors produced by transformed cells such as insulin-like growth factor (IGF)<sup>11,12</sup>, fibroblast-derived growth factor<sup>13,14</sup> and the transforming growth factors (TGFs)<sup>15-20</sup>, may also act as autocrine regulators of proliferation. Besides the specificity mediated by regulation of the production of growth factors, cellular specificity could also be controlled at several other levels—the most obvious being by binding of ligand to specific receptors present only on target cells. In addition the binding of one growth factor to its specific receptor can also alter the affinity of another growth factor for its receptor (for example, PDGF and the EGF receptor<sup>21,22</sup>). Conversely two growth factors may, as appears to be the case with  $\alpha$ TGFs and EGF, bind to the same receptor<sup>23,24</sup>.

It is clear that binding of different growth factors to their specific receptors can induce a cascade of biochemical events including rapid changes in ion movements and intracellular pH, stimulation of tyrosine specific protein kinases and several other changes which can culminate in DNA synthesis and proliferation

of certain target cells<sup>1,4-6</sup>. It seems likely that at least in the case of the EGF receptor the primary function of EGF may be to induce cross-linking or conformational changes of receptors, and that following such an activation step, all the 'information' necessary for triggering a proliferative response may reside in the receptor itself (see ref. 1). One known function intrinsic to the EGF receptor is its ability to phosphorylate tyrosine residues<sup>25-28</sup>, a property shared with five of the putative transforming proteins of the family of retroviruses whose oncogenes are structurally related to *src* but not by two others, the proteins encoded by *mos* and *erb-B*<sup>29</sup>. At present this tyrosine kinase activity provides the only functional activity associated with the oncogenes of this subset of retroviruses.

Here we report amino acid sequence analysis of six distinct peptides from human EGF receptors isolated by monoclonal immunoaffinity purification from A431 cells and placenta, and show that 74 out of 83 of the residues sequenced are identical to those of the transforming protein encoded by the *v-erb-B* oncogene of avian erythroblastosis virus (AEV)<sup>30</sup>. The amino acid sequences of several other peptides purified from the EGF receptor could not be aligned with sequences of the *v-erb-B* protein suggesting that AEV has acquired cellular sequences encoding only a portion of the avian EGF receptor. Several lines of evidence suggest that the *v-erb-B* oncogene encodes only the transmembrane region of the EGF receptor and the domain associated with the tyrosine kinase activity. These results suggest a hypothesis that the *src* related subset of oncogenes, which includes *v-erb-B*, may be derived from cellular sequences which encode growth factor receptors and produce transformation through expression of uncontrolled receptor functions.

## Purification of EGF receptor

EGF receptors can be detected in a variety of cells either by measurement of EGF binding (reviewed in ref. 31), by cross-linking of labelled EGF to its receptor (reviewed in ref. 32) or by using monoclonal antibodies<sup>33-38</sup>. In this study the receptor has been purified from two sources: the human epidermoid carcinoma cell line A431, which expresses about 50 times more receptors than the majority of other cells<sup>39,40</sup>, and human placenta<sup>41</sup> which is a readily available normal tissue. The recent isolation of monoclonal antibodies which recognize the human EGF receptor<sup>34,38</sup> has made it possible to use immunoaffinity chromatography for receptor purification. Here we compare by peptide mapping EGF receptor protein purified by either immunoaffinity or EGF affinity chromatography<sup>26</sup> and also compare the structures of the A431 and placental receptors.

A radioimmunoassay (RIA) which uses a monoclonal antibody (R1) has been used to quantitate various preparative

§ Present address: Department of Biochemistry, University of Otago, Dunedin, New Zealand.

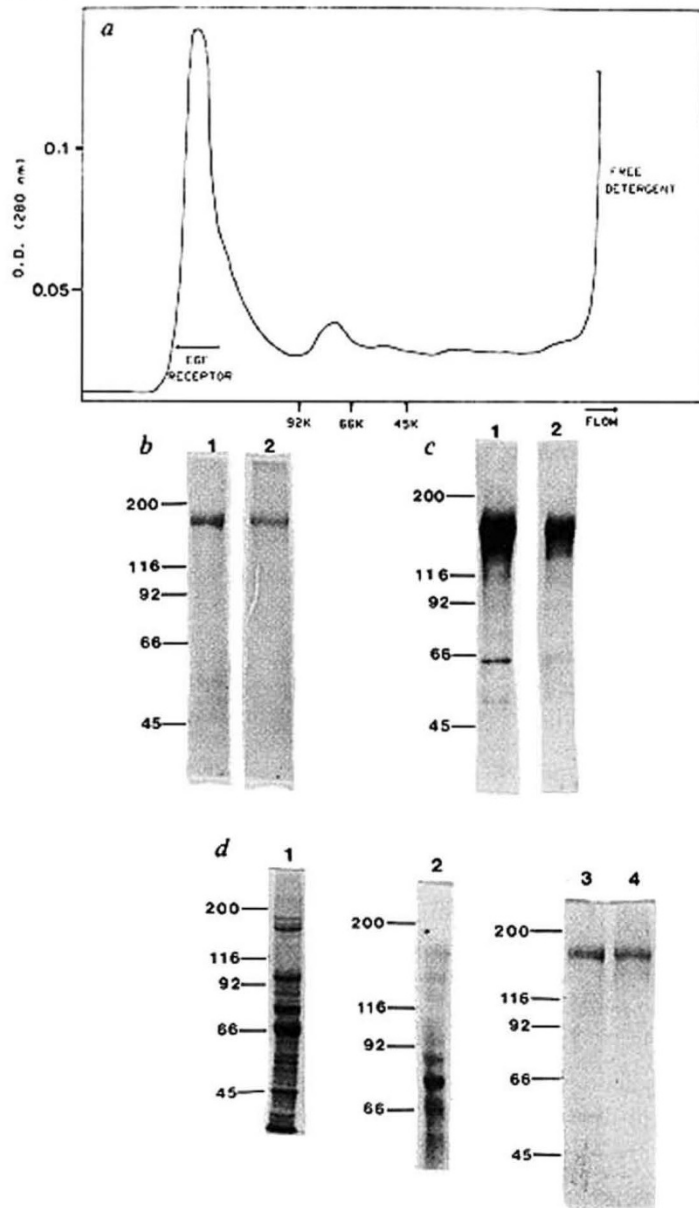


**Fig. 1** Immunopurification of EGF receptor from A431 cells and human placenta. **a**, Purification by gel permeation. Reduced and alkylated receptor was purified on a TSK4000 column (0.7×60 cm, LKB) using 0.1 M potassium dihydrogen phosphate buffer pH 4.5 containing 6 M guanidine HCl at a flow rate of 0.5 ml min<sup>-1</sup> (ref. 43). The absorbance of the eluate was monitored at 280 nm (molecular weights of protein standards are indicated) and 0.25 ml fractions were collected. Fractions containing EGF receptor were dialysed against 10 mM ammonium bicarbonate. Panels **b–d** show 7% polyacrylamide SDS gels<sup>72</sup> used to monitor purification (MW×10<sup>-3</sup> of protein standards are indicated). **b**, R1 purified A431 EGF receptor: track 1, pH 3 eluate from R1 immunoaffinity matrix; track 2, eluate from the TSK4000 column. **c**, 29-1 purified A431 EGF receptor: track 1, pH 3 eluate from 29-1 immunoaffinity matrix; track 2, eluate from SDS preparative gel electrophoresis. **d**, Lectin and R1 purified placental EGF receptor: track 1, placental vesicles; track 2, eluate from lectin affinity matrix; track 3, eluate from R1 immunoaffinity matrix; track 4, eluate from the TSK4000 column.

**Methods:** Approximately 2×10<sup>9</sup> A431 cells were washed in calcium and magnesium free phosphate-buffered saline (PBS) and solubilized in 400 ml lysis buffer (50 mM Tris-HCl pH 7.4, 0.15 M NaCl, 5 mM EGTA, 0.1% bovine serum albumin, 1% NP40, 25 mM benzamide, 0.2 mM PMSF, 10 µg ml<sup>-1</sup> leupeptin). After filtration through muslin the lysate was adjusted to pH 8.5 and centrifuged at 100,000g<sub>max</sub> for 30 min. The supernatant was incubated for 2 h at 4°C with immunoaffinity matrix, which consisted of 15 mg of monoclonal antibody R1<sup>34</sup> coupled to 15 ml of Affi-Gel 10 (BioRad). Unbound lysate was removed by suction through a 0.4 micron filter. The matrix was then washed by gentle agitation and filtration with 500 ml PBS, containing 0.65 M NaCl and 0.1% NP40, followed by 500 ml PBS, containing 0.1% NP40. The EGF receptor was eluted by gentle agitation and filtration of the matrix with 2×10 ml aliquots of 50 mM sodium citrate pH 3, containing 0.05% NP40 for 10 min each. Eluates were adjusted to pH 7. The yield of receptor was approximately 250 µg, measured by Bradford technique<sup>70</sup> or by amino acid analysis after gel permeation HPLC (see below). Alternatively EGF receptor was purified from A431 cells using monoclonal antibody 29-1<sup>38</sup> coupled to CNBr-activated Sepharose (Pharmacia) at 5–10 mg ml<sup>-1</sup>. The purification procedure used was similar to that described for the R1 immunoaffinity matrix except that the EGF receptor was phosphorylated whilst bound to the matrix with 50 µCi [γ-<sup>32</sup>P]ATP (3,000 Ci mmol<sup>-1</sup>, Amersham International) in the presence of 3 mM MnCl<sub>2</sub>. Eluted receptor was further purified by preparative SDS gel electrophoresis, followed by dialysis against 10% methanol at 4°C. For purification of placental EGF receptor, vesicles were made from syncytiotrophoblast microvilli by a modification of the method of Smith *et al.*<sup>71</sup>, using 2 mM EGTA in all buffers. Vesicles were solubilized by addition of an equal volume of 100 mM HEPES pH 7.4, 0.15 M NaCl and 5% Triton X-100. After centrifuging at 100,000g<sub>max</sub> for 30 min the supernatant was incubated for 1 h at room temperature with 200 mg wheat germ agglutinin coupled to 10 g of Affi-Gel (BioRad). The lectin matrix was washed by filtration through a 0.4 µm filter with 100 ml of PBS, containing 0.1% Triton X-100. Bound protein was eluted by agitation and filtration with 2×15 ml aliquots of 0.25 M *N*-acetylglucosamine in 10 mM HEPES pH 7.4, 0.1% Triton X-100 for 15 min each. The eluate was incubated with R1 immuno-affinity matrix (15 mg antibody per placenta) for 2 h at 20°C, followed by extensive washing and elution as described above for receptor purification from A431 cells. The yield of EGF receptor per placenta was 25 µg, measured by Bradford technique<sup>70</sup> and amino acid analysis after gel permeation HPLC (see below). Solutions containing EGF receptor were lyophilized and resuspended in 0.5 M Tris-HCl pH 8.5, 6 M guanidine hydrochloride (Schwarz-Mann) at 0.5–1 mg ml<sup>-1</sup>. After incubation at 37°C for 16 h with 10 mM dithiothreitol, cysteine residues were alkylated with <sup>14</sup>C-iodoacetamide (40–60 mCi mmol<sup>-1</sup>, Amersham International) as described previously<sup>44</sup>.

techniques<sup>42</sup>. Receptors from A431 cells and placenta were both found to be unstable in detergent-solubilized whole cell or tissue lysates, perhaps as a result of the release of proteases from the cellular lysosomal compartment. This problem was overcome for placenta by the preparation of syncytiotrophoblast microvilli plasma membranes and as a result a 50-fold purification with a 30% yield of receptor was achieved. Unfortunately, with A431 cells the yield of receptor in plasma membrane preparations was impractically low. However, quantitative studies with the receptor RIA showed that rapid adjustment of the lysate pH to 8.5 followed by fast immunoaffinity chromatography of whole cell lysates minimised the effects of the proteases.

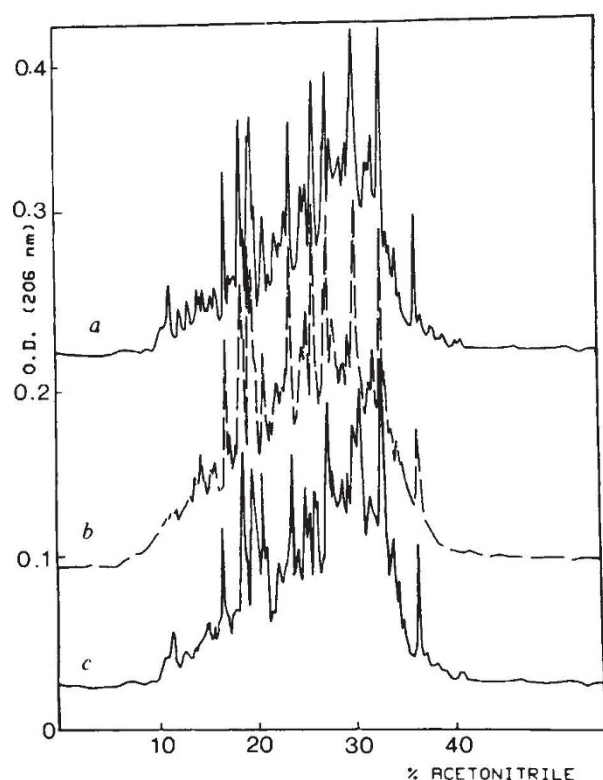
Placental membranes were solubilized and glycoproteins separated by wheat germ agglutinin (WGA) affinity chromatography to achieve a partial purification. EGF receptor was then purified from the placental glycoprotein fraction or from A431



cell lysates by immunoaffinity chromatography on either monoclonal antibodies R1<sup>34</sup> or 29-1<sup>38</sup> immobilised on agarose or Sepharose respectively. Nonspecifically bound protein was removed by washing the columns with a high salt buffer and the receptor was eluted at pH 3. The receptors were further purified either on preparative SDS-polyacrylamide gels or by gel permeation HPLC in guanidine solutions<sup>43</sup>. Details of the methods used and the yields of purified receptors are given in the legend to Fig. 1. Since the EGF binding and protein kinase activity were partially destroyed during purification, receptor was also purified by EGF affinity chromatography<sup>26</sup>. Comparative HPLC tryptic peptide maps were then carried out to establish the purity and structural similarity of receptor prepared by immunoaffinity chromatography from A431 cells and placental tissue. The peptide maps of the receptors (Fig. 2) showed that the elution profiles of the receptor tryptic peptides were very



similar whether receptor was purified by EGF affinity or by immunoaffinity chromatography; from A431 cells or from placental tissue.



**Fig. 2** Reverse phase HPLC analysis of tryptic peptides from EGF receptor purified by three different affinity methods. *a*, Immunoaffinity purified A431 receptor. EGF receptor was purified from A431 cells using the monoclonal antibody R1 and gel permeation chromatography in guanidine solutions as described in Fig. 1. Pooled fractions containing the 175,000 MW EGF receptor were dialysed against 10 mM ammonium bicarbonate, lyophilized and resuspended in 500  $\mu$ l of 100 mM ammonium bicarbonate, 10 mM  $\text{CaCl}_2$ . TPCK treated trypsin (Sigma) was then added (100:1, receptor:trypsin, w/w). This mixture was incubated at 37 °C for 12 h, a further identical aliquot of trypsin added and the incubation continued for another 12 h. Peptides were then loaded directly onto a Synchropak RPP C18 reverse phase HPLC column (Synchrom, Linden, Indiana, 4.6  $\times$  75 mm) equilibrated in 0.1% trifluoroacetic acid (TFA, Rathburn, Scotland) and a gradient from zero to 45% acetonitrile was run over 45 min at 1 ml min<sup>-1</sup> and 1 ml fractions collected<sup>46</sup>. A Waters HPLC system including two M6000 A pumps, a U6K manual injector and a 660 solvent programmer with 2 LKB 2138 Uvicord S absorbance detectors with filters at 206 nm and 280 nm was used<sup>45</sup>. The figure shows the optical density of the eluate at 206 nm plotted against acetonitrile concentration. *b*, EGF affinity purified receptor. EGF receptor was purified from A431 cells using wheat germ agglutinin affinity chromatography followed by EGF affinity chromatography<sup>26</sup>. A431 cells were lysed as described in the legend to Fig. 1. Purification of this lysate on the WGA affinity column was identical to that described for the placental preparation (Fig. 1). The eluate from the WGA affinity column was mixed with 5 ml of Affi-Gel 10 (BioRad) having 1 mg of bound EGF. The mixture was tumbled for 4 h at room temperature before washing the immobilized EGF receptor with 100 ml of PBS, 0.1% Triton X-100. Receptor was then eluted with 10 mM ethanolamine, pH 9.7, 0.1% Triton X-100. This eluate was further purified on a TSK4000 gel permeation column as described in Fig. 1. Fractions containing EGF receptor were pooled, dialysed, lyophilized and trypsinized as described for *a* above. The resulting tryptic peptides were separated by reverse phase HPLC under identical conditions to those described above. *c*, Immunoaffinity purified placental receptor. EGF receptor was purified from fresh term human placenta as described in the legend to Fig. 1. The receptor was digested with trypsin and peptides separated by reverse phase HPLC as described above. Peptides in the column breakthrough are not shown.

## Amino acid sequence

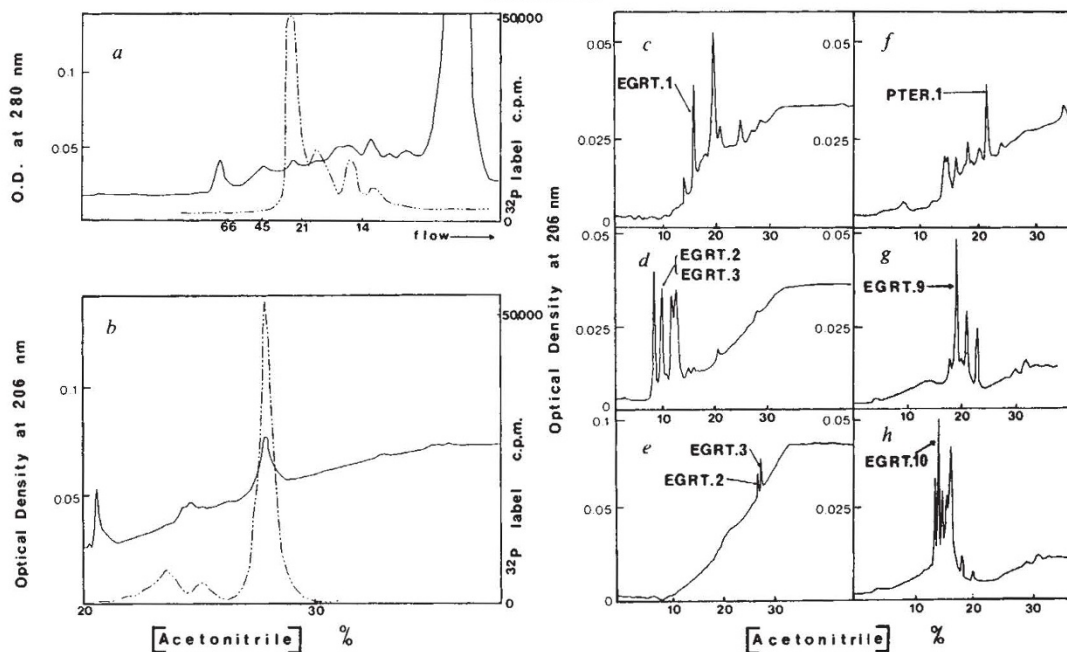
Receptor was purified by immunoaffinity chromatography followed by either preparative SDS gel electrophoresis or by gel permeation HPLC in guanidine<sup>43</sup> after reduction and alkylation<sup>44</sup> to cleave disulphide bonds (see Fig. 1). Purified receptor was then digested with trypsin or cyanogen bromide (see Figs 2 and 3) and peptides were separated by preparative reverse phase HPLC<sup>45,46</sup> (Fig. 3). Amino acid sequences were determined with a gas phase sequencer constructed and operated as described by Hewick *et al.*<sup>47</sup>, using the analytical techniques for the quantitation of phenyl thiohydantoin (PTH) amino acids described by M.D.W. *et al.*<sup>48</sup>. The quantitative data for analysis of six peptides are shown in Fig. 4.

The amino acid sequences of 14 different peptides from the human EGF receptor, three from placenta and 11 from A431 cells, were compared with sequences in an oncogene sequence data base (set up at ICRF using published sequences) by the rapid search techniques of Wilbur and Lipman<sup>49</sup>. A remarkable identity was found between the sequences of six of these peptides and regions of the predicted sequence of the putative transforming protein *v-erb-B* of the AEV-H isolate of avian erythroblastosis virus<sup>30</sup>. Of the 83 amino acid residues from these six sequenced peptides, 74 residues were identical and four showed conservative substitutions when they were aligned with the *v-erb-B* encoded protein sequence, as shown in Fig. 5. Peptide 1 was located near the amino terminus of the *v-erb-B* protein (residues 107–125) and peptide 6 at the carboxy terminus (residues 583–599), with the other four peptides in between. It was not necessary to introduce any deletions or insertions into the sequence to optimize the alignments.

Although the full extent of the similarity between the *v-erb-B* protein and EGF receptor sequences is not revealed by these limited sequence studies, it is likely that the region of the *v-erb-B* protein from residue 107 to the carboxy terminus has extensive homology to the EGF receptor. The degree of identity observed is very high and since the *v-erb-B* sequences of AEV were presumably of avian origin<sup>30</sup>, while the EGF receptor sequences were from the human protein, it is likely that the *v-erb-B* sequences were mainly acquired by AEV from those cellular sequences which encode the avian EGF receptor. This suggests that the *c-erb-B* locus encodes the EGF receptor in humans and birds.

The amino acid sequence of 8 of the 14 peptides purified from the EGF receptor (data not shown) could not be aligned with the predicted sequences of the *v-erb-B* protein. Since the polypeptide backbone of the EGF receptor glycoprotein is thought to be about 1,250 amino acids<sup>50</sup> and the predicted *v-erb-B* protein is only 604 amino acids<sup>30</sup> the most likely explanation is that these eight peptides are encoded by a region of *c-erb-B* which has not been acquired by AEV. This could have arisen by a recombination event(s) which resulted in only a part of the EGF receptor coding sequences being acquired by AEV. Although it is possible that DNA rearrangements of receptor coding sequences occur similar to those found with immunoglobulins, it is more likely that differential mRNA splicing would be involved in any such recombination events. It has been shown that avian cells contain two *c-erb-B* related transcripts<sup>51</sup> and studies of the biosynthesis of the EGF receptor in A431 cells suggest that both a normal and a truncated receptor may be synthesized<sup>50</sup>. Alternatively two or more loci encoding polypeptides having very similar amino acid sequences to those of the EGF receptor exist on chromosome 7 (see below). An example of two closely related putative transforming proteins with tyrosine kinase activity has been reported in studies of the avian retroviruses Rous sarcoma virus (RSV) and Yamaguchi Y63<sup>52</sup>. The predicted amino acid sequences of the proteins encoded by the *src* and *yes* oncogenes were shown to have 82% match over a region covering 436 amino acid residues (while the DNA sequences showed only 31% overall match) and presumably the chicken genome contains both *src* and *yes* proto-oncogenes encoding separate proteins sharing extensive regions of sequence. It is not known whether human *c-src* and *c-yes*





**Fig. 3** Purification of peptides from EGF receptor for sequence analysis. **a**, Cyanogen bromide cleavage and fractionation of peptides by size.  $^{32}\text{P}$ -labelled EGF receptor in ammonium bicarbonate solution was lyophilized and resuspended in 70% formic acid. Cyanogen bromide was added under nitrogen, the tube sealed and incubated in the dark at room temperature for 24 h. Formic acid and excess cyanogen bromide were removed by repeated cycles of drying and resuspension in water using a Speed-vac concentrator (Savant). The dry sample was resuspended in 0.1 M potassium dihydrogen phosphate buffer, pH 4.5 containing 6 M guanidine HCl and the peptides separated by gel permeation HPLC on a TSK3000 column ( $0.7 \times 60$  cm, LKB) equilibrated in the same buffer at a flow rate of  $0.3 \text{ ml min}^{-1}$ . The optical density of the eluate was monitored at 280 nm (---) and 0.3 ml fractions collected and counted for  $^{32}\text{P}$  (—). Molecular weights  $\times 10^{-3}$  of protein standards are indicated. **b**, Subfractionation of cyanogen bromide fragments. The peak from the TSK3000 column containing most of the  $^{32}\text{P}$ -label was pooled and dialysed against 10 mM ammonium bicarbonate. After lyophilization, the sample was redissolved in 0.1% TFA and peptides separated by reverse phase HPLC on a Synchropak RPP C18 column (see Fig. 2) equilibrated in 0.1% TFA, 10% acetonitrile<sup>45,46</sup>. A gradient of 10–40% acetonitrile run over 60 min was used to elute peptides, at a flow rate of  $1 \text{ ml min}^{-1}$ . The optical density of the eluate was monitored at 206 nm (---) and 1 ml fractions were collected and counted for  $^{32}\text{P}$ -label (—). **c**, The fractions corresponding to 23–24% acetonitrile from the HPLC analysis of A431 EGF receptor tryptic peptides were pooled. Peptides were further purified by reverse-phase HPLC on a Synchropak RPP C18 column equilibrated in 10 mM ammonium acetate buffer pH 6.5. A linear gradient of 0–45% acetonitrile was run over 45 min at a flow rate of  $1 \text{ ml min}^{-1}$ . The optical density of the eluate was monitored at 206 nm and 0.5 ml fractions collected. **d**, The fractions corresponding to 19–20% acetonitrile from the reverse-phase HPLC purification of A431 EGF receptor tryptic peptides (Fig. 2a) were pooled. Peptides were separated as described in **c**. **e**, The peak fractions arrowed in **d** were pooled and peptides subfractionated by reverse-phase HPLC on a  $\mu$  Bondapak phenyl column ( $0.46 \times 25$  cm, Waters Assoc.) equilibrated in 0.1% TFA. A linear gradient of 0–45% acetonitrile over 45 min was used to elute peptides, at a flow rate of  $1 \text{ ml min}^{-1}$ . The optical density of the eluate was monitored at 206 nm and 0.2 ml fractions collected. **f**, The fractions corresponding to 27–28% acetonitrile concentration from the reverse-phase HPLC analysis of placental EGF receptor tryptic peptides (Fig. 2c) were pooled. Peptides were subfractionated as described in **c**. **g**, The fractions corresponding to 25–26% acetonitrile from the reverse-phase HPLC purification of A431 EGF receptor tryptic peptides (Fig. 2a) were pooled. Peptides were subfractionated as described in **c**. **h**, The fractions corresponding to 21–22% acetonitrile from the reverse-phase HPLC analysis of A431 EGF receptor tryptic peptides (Fig. 2a) were pooled. Peptides were subfractionated as described in **c**.

are encoded by closely linked loci. However, analysis of human-mouse somatic cell hybrids has shown that the locus encoding the human EGF receptor is on chromosome 7 (7p13-7q22)<sup>53–55</sup> and that for *c-erb-B* is in the same region of this chromosome (7pter-7q22)<sup>56</sup>. This observation supports the concept that the *c-erb-B* locus encodes the human EGF receptor or is very closely linked to that encoding the EGF receptor.

Analysis of EGF receptor mRNA transcripts found in A431 and avian cells by cDNA cloning is currently in progress which, together with analysis of human and avian genomic clones, should define the relationship of *c-erb-B* and the locus encoding the EGF receptor (A.U. *et al.*, in preparation).

### Shared regions of sequence

Several lines of evidence suggest that the EGF receptor protein can be divided into three major domains; an EGF binding domain which lies external to the plasma membrane, a trans-membrane domain and a cytoplasmic kinase domain having both the kinase activity and the autophosphorylation sites.

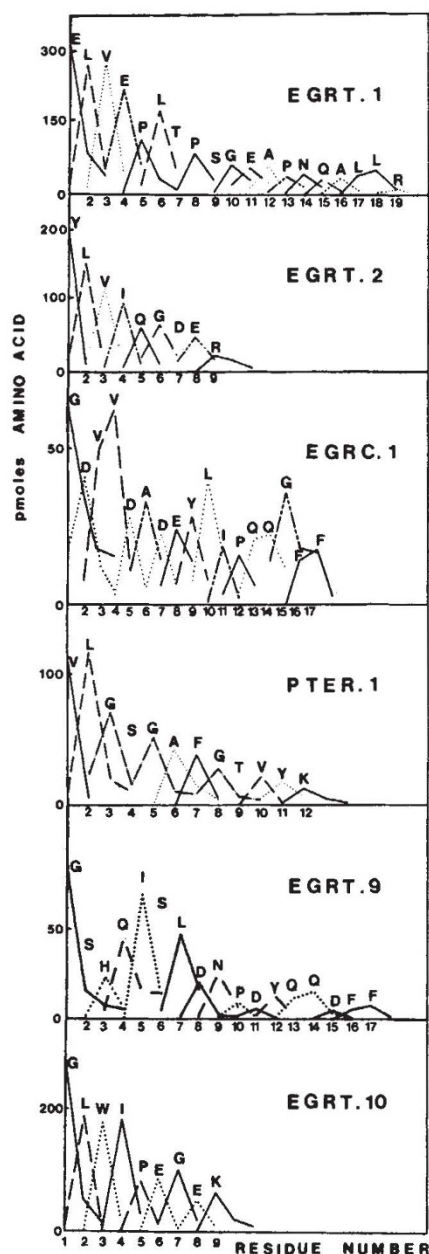
Investigations of receptor biosynthesis show that the A431 receptor is a glycoprotein of apparent molecular weight (MW) 175,000, having  $\sim 37,000$  MW of oligosaccharide side chains with a polypeptide backbone of  $\sim 138,000$  MW ( $\sim 1,250$  amino acids). Limited proteolysis of the mature receptor suggests that the domain external to the plasma membrane which contains

the oligosaccharide side chains and the antigenic sites for monoclonal R1 has a MW of  $\sim 115,000$  ( $\sim 710$  amino acids)<sup>50</sup>. Several studies show that the EGF binding site is external to the plasma membrane<sup>25,31,32</sup>.

The location of the tyrosine kinase enzymatic activity and the autophosphorylation sites on the cytoplasmic domain is supported by studies made using A431 and placental membrane vesicles (J.D. and M.D.W., in preparation). These show that EGF stimulated tyrosine kinase activity directed towards artificial substrates or towards autophosphorylation sites is significantly activated only after membrane permeabilization. Furthermore, the tyrosine kinase activity can phosphorylate pp36—a protein known to be located at the cytoplasmic side of the membrane<sup>57</sup>. In addition recent studies show that antibodies raised against synthetic peptides from pp60<sup>v-src</sup> recognize antigenic sites on the human EGF receptor that are from regions of sequence homologous to the sequence of *v-erb-B* (see below). These sites are only accessible in permeabilized cells (J.S. *et al.*, in preparation).

Autophosphorylation sites are located within peptide 5 (EGRC.1), a 20,000 MW cyanogen bromide fragment which contains 70% of the  $^{32}\text{P}$ -label present in the autophosphorylated receptor (see Fig. 3a). Although the precise location of the residues phosphorylated has not been determined a consensus tyrosine phosphorylation sequence<sup>58</sup> was found near the amino





**Fig. 4** Sequence analysis of peptides from the EGF receptor. Peptides were purified as described in Fig. 3. Sequence determination of each peptide was carried out using a gas phase sequencer assembled and operated as described by Hewick *et al.*<sup>47</sup>. PTH amino acids were analysed by HPLC using a Zorbax C8 column (4.6 × 150 mm, Dupont) at 43 °C with a linear gradient over 8 min of acetonitrile from 24% to 38% at a flow rate of 2 ml min<sup>-1</sup> using 9 mM sodium acetate buffer pH 4.1<sup>48</sup>. A Waters HPLC system including two M6000 A pumps, a WISP autoinjector and system controller with a Beckman Model 160 detector was used. The recovery of PTH amino acids at each degradative cycle was measured using an integrative recorder (Waters Data module). The amounts of each peptide analysed (pmol) were measured by amino acid analysis (EGRT.1, 380; EGRT.2, 260; EGRC.1, 80; PTER.1, 145; EGRT.9, 100; EGRT.10, 325). The analysis for serine and threonine could not be accurately measured due to the presence of multiple peaks obtained during analysis of the PTH amino acids. The presence of these amino acids is thus indicated without quantitative data; these residues are assigned to the sequence using semi-quantitative recovery data based on peak heights rather than areas. Prior to loading peptides, fibre glass disks were treated with polybrene and glycylglycine and precycled for ten cycles. Each peptide was sequenced twice; on the second run of peptide EGRC.1 filters were treated with polybrene and cysteic acid and precycled ten times to clarify the assignment of an amino terminal glycine residue. However the background glycine at step 1 is still significant and this residue may be incorrect.

terminus of this peptide. Therefore we believe that the tyrosine phosphorylation sites lie within the cytoplasmic domain of the EGF receptor which is contained in the sequence shared with the *v-erb-B* protein.

Preliminary nucleotide sequence analysis (A.U. *et al.*, in preparation) of cDNA clones selected from a placental cDNA library using synthetic oligonucleotide probes synthesized on the basis of the receptor amino acid sequence shows that the predicted carboxyl terminus of the EGF receptor extends 32 amino acids from an amino acid equivalent to residue 601 (see Fig. 5) of the predicted *v-erb-B* protein sequence. This analysis when complete will show the precise size and sequence of the presumptive cytoplasmic domain of the EGF receptor. The approximate molecular weight of this domain would be 60,000 (or 545 amino acids) since that part which is external to the membrane is thought to have a molecular weight of 115,000 (see above)<sup>50</sup>. Thus the cytoplasmic domain would be predicted to be similar in size to that region of the *v-erb-B* protein which is carboxy terminal to a putative transmembrane sequence (see Fig. 5 and ref. 30) at residues 66–88. This carboxy terminal region of *v-erb-B* would have a molecular weight of 56,760 and would contain 516 amino acids.

The putative transmembrane sequence of the *v-erb-B* protein is not preceded by a signal sequence for membrane insertion. Nevertheless, immunofluorescence studies of AEV transformed cells show that the *v-erb-B* protein has antigenic sites external to the plasma membrane<sup>59</sup>. This external region probably corresponds to the 65 residue amino terminal section that precedes the putative transmembrane sequence and contains three asparagine residues which have the oligosaccharide attachment recognition sequence Asn-X-Ser or Thr. Some or all of these residues may be glycosylated since *in vitro* translation studies of mRNA from AEV infected cells show that post translational processing of nascent polypeptides occurs in the presence of membrane vesicles<sup>59,60</sup>.

Together these studies suggest that the predicted *v-erb-B* transforming protein closely resembles the transmembrane region of the EGF receptor and the domain which is thought to be cytoplasmic. If the *v-erb-B* sequence was acquired from the gene encoding the EGF receptor then the *v-erb-B* protein represents a truncated receptor which lacks the EGF binding domain. It is particularly interesting that studies of EGF receptor biosynthesis in A431 cells have suggested that a polypeptide equivalent to the external domain of the receptor (of MW 115,000) is synthesized<sup>50</sup> in addition to the normal receptor. Further studies are necessary to understand the origin of this truncated receptor but the results show that defective receptors may be synthesized by this human tumour cell line.

## Growth factors and transformation

Recently it has been shown that the transforming protein of simian sarcoma virus has a close structural and functional relationship to the growth factor PDGF<sup>7–10</sup> supporting the hypothesis that autocrine growth factor production may be involved in abnormal growth control and neoplasia. These observations together with those presented here illustrate two distinct but related mechanisms for subversion of normal growth regulation. In the case of SSV the oncogene encodes a growth factor which can act as a mitogen for target cells having PDGF receptors<sup>10</sup>. AEV on the other hand appears to utilize a different mechanism where a part of a growth factor receptor which is thought to be involved in transducing the EGF signal may be expressed in transformed cells. The absence of the EGF binding domain might remove the control generated by ligand binding and the result could be the continuous generation of a signal equivalent to that produced by EGF, causing cells to proliferate rapidly. How this could result in the block in differentiation observed in AEV infected haemopoietic cells<sup>61</sup> is unclear. However, EGF has been shown to promote proliferation while inhibiting terminal differentiation of human keratinocytes<sup>31</sup>.

The ES4 strain of AEV has two oncogenes *v-erb-A* and *v-erb-B*, which are thought to encode proteins of MWs 75,000



*Src* 1 MGSSKSKPKDPSQRHSLPEPDSHGGFASQSTPDETAAPDAHRNPSRS

*Src* 51 FGTVATEPKLFWGNTSDTVTSQRAGALAGGVTTFFVALYDYESHTETDL

*Src* 101 SFKKGERLQIVNTEGDWHLAHLSTTGOTGYIPSNYVAPSDSIQREWHYF  
MKCAHFIDGPHCVKA

*Src* 151 GKITRRESERLLNPNPGRGTFLVRKSETAKGAYCLSVSDFNAGKPNVK  
*Erb-B* 16 CPAGVLGENDTLVRKYADANAVCQLCHPNTCRGCKGPGLEGCPNGSKTFS

*Src* 201 HYKIYKLYSGGFYITSRTOFGSLQDLVAYYSKHAQGLCHRLANVCPTSKP  
*Erb-B* 66 IAAGVVGGGLLCLVVGGLGIGLYLRRRHIVYKRTLRLLOERLEVEPTLP  
ELVEPTLP

*Src* 251 QTGLAKDAHCIPRESLRLEAK-LGGGCFGEVNMQTN--D---TTTVAI  
*Erb-B* 115 SGARPNGAHLRIKTEFKKVKVLGGGAFDTIYKGLHIFEGEKYKIPVAI  
SGARPNGALL--VLGGGAFDTIYKGLHIFEGEK

*Src* 295 KILKPGTMSPEA---FLOEQGVMMKSLHEKLVLYAVVSEETIYIVICV  
*Erb-B* 163 KELREAT-SPKANKELIDERYMVKSLVDNPHVGLIGLITLTVGLITGLH

*Src* 346 SKGSLDPLFKGEMCKYLRLPQ-LVDMARQIASQMAVYV-AMNYVHDLRA  
*Erb-B* 214 PYGCLLDY-IRE-HKNDIGSGYLLNLCVGIKGMHNYLEER-LVHDL-A

*Src* 390 A-NILVGENLYCKVADFLARLIE-D-NEYTARQAKFKIKHTAPEARLY  
*Erb-B* 260 ARNVLVKTPGHYKITDPLAKLLGABEKYHREGS-KYPIKMMALLESILH

*Src* 437 GRFTIKSDVMSFGILLTELTKGRVYPMVNVRELDQVERGYMPCPPE  
*Erb-B* 389 RIYTHGSDVMSYGVTVHMLTFSKPYDGIAPSEISSVLEKGERLPGPPI

*Src* 487 CPESLHLMCOCHRKDPPEERFTKYLRQLLPACVLEVAE  
*Erb-B* 359 CTIDVYIMVYKCMHIDADSAPKPRELIRAFSKMARDPPRYLVIGQDERMH  
\*LVIGQDERMH

*Erb-B* 489 LPSPDTSKFYRTLMEEDMEIDVDADEYLVPFGGFFNSPSTSRTPLLSSL  
\*EDVDADEYLVPFGGFF

*Erb-B* 459 SATSNNSATNCIDRNGQGHVREDSDVQRYSSDPTGNFLEESIDGFLPA

*Erb-B* 589 PEYVNLMPKKPSTAMVQNGIYNFISLTAISKLPMSRYQNSHSTAVDNP

*Erb-B* 559 EYLNTNQSPAKTVFESSPYWISQGNHNLNLPFYGGFFLFTSCS  
\*SHHSLNLPFYGGFF

**Fig. 5** The relationship between the amino acid sequences of the EGF receptor peptides and the predicted amino acid sequences of the putative transforming proteins of *v-src* and *v-erb-B*. The predicted amino acid sequence of the *v-src* gene product (pp60<sup>v-src</sup>) is translated from the presumptive initiation codon at nucleotide 7,129 of the Prague C strain of Rous sarcoma virus<sup>73</sup>. The predicted amino acid sequence of the *v-erb-B* gene product is translated from the presumptive initiation codon at nucleotide 155 of the *v-erb-B* gene in AEV-H<sup>30</sup>. The partial amino acid sequences of the six peptides purified from the EGF receptor are shown (underlined): 1, EGRT.1; 2, PTER.1; 3, EGRT.10; 4, EGRT.2; 5, EGRC.1; 6, EGRT.9. Letters in bold type represent residues shared by pp60<sup>v-src</sup> and the *v-erb-B* protein. Residues shared by the EGF receptor peptides and *v-erb-B* protein or pp60<sup>v-src</sup> are in bold type. ● indicates amino acid residues, which are common to the putative transforming proteins of *v-erb-B*, *v-src*, *v-fes*, *v-fps*, *v-yes* and *v-abl*. †, Phosphoacceptor tyrosine of pp60<sup>v-src</sup><sup>74</sup>. ▲, Possible N-linked glycosylation sites at the amino terminus of the *v-erb-B* protein. — indicates the putative transmembrane sequence in the *v-erb-B* protein. \*, Amino acid residues which would be expected to produce enzymatic or cyanogen bromide cleavages to generate the observed peptides. Numbers to the left of the sequences are residue numbers taking the presumptive initiation methionine as 1 in both cases. Sequences were aligned using a computer program<sup>75</sup> to optimize match.

and 65,000 respectively (for a recent review see ref. 62). Cells transformed by AEV *in vitro* and *in vivo* have the properties of erythroblasts which are late erythroid progenitors, although the target cells themselves may be earlier erythroid precursor cells. AEV can also transform fibroblasts and induce sarcomas. Evidence from deletion mutants<sup>63</sup> and from an isolate (AEV-H<sup>30</sup>) which lacks the *v-erb-A* gene suggests that the *v-erb-B* gene alone can induce transformation. This is supported by studies which show that RAV-1, a leucosis virus which has no oncogene, can activate the *c-erb-B* gene by a promoter insertion mechanism<sup>64</sup> perhaps similar to that presently being unravelled for *c-myc* activation<sup>65-67</sup>. It is possible that RAV-1 could induce expression of a normal receptor or a truncated receptor. EGF receptors have not generally been detected on haemopoietic cells by EGF binding studies but since these studies are limited in scope and sensitivity a more rigorous survey is needed before conclusions about normal receptor expression in different haemopoietic cell types can be made. Although many normal

cells express 10–100,000 EGF receptors<sup>31</sup> only very low levels of *c-erb-B* transcripts have been found in normal chicken fibroblasts<sup>51</sup>. However a recent study of normal and neoplastic human lymphocytes suggests that both types of cells contain *c-erb-B* related transcripts. Clearly more detailed studies will be necessary to search for normal or altered *c-erb-B* or *v-erb-B* transcripts in normal, neoplastic or AEV transformed cells.

Previous reports have shown that the predicted amino acid sequences of the putative viral transforming proteins encoded by the oncogenes *erb-B*, *src*, *yes*, *fes*, *fps*, *mos* and *abl* show regions of similarity, implying homology<sup>29,30,69</sup>. In the case of *src*, *yes*, *fes*, *fps* and *abl* the putative transforming proteins have been shown to have tyrosine kinase activity (reviewed in ref. 29) but as yet those encoded by *erb-B*<sup>30</sup> and *mos* have not. Since the receptors for EGF and  $\alpha$ TGF, PDGF, insulin and IGF-I also have associated tyrosine kinases, the structural relationship between the *v-erb-B* transforming protein and the EGF receptor observed here, suggests that other oncogenes from this subset of retroviruses could be derived in part from sequences encoding these or other growth factor receptors. Alternatively the other oncogenes in this group may be derived from sequences which encode proteins having a less direct functional relationship to the EGF and other receptors. Further study will clarify the precise function of growth factors, their receptors and *src* related oncogenes in normal and malignant growth.

M.D.W. thanks B. Ozanne and E. Rozengurt for introducing him to the field of study involving growth factors and their receptors. We acknowledge P. Parker, P. Stroobant, H. S. Earp and M. Fried for helpful discussions and advice, W. Gullick for help with receptor quantitation techniques and P. Bennett and R. Philp for technical assistance. We also thank W. J. Wilbur and D. J. Lipman for computer programs and N. Teich and the ICRF Computer Unit for their help. M.D.W. thanks S. J. for continued support. We are indebted to Queen Charlotte's Hospital and to E. J. Owen for help in providing placental tissue. This work was supported in part by grants from the USNIH (CA 25820) (J.S.) and from the US-Israel Binational Science Foundation (J.S.).

Received 18 January; accepted 23 January 1984.

- Guroff, G. (ed.) *Growth and Maturation Factors* (Wiley, New York, 1983).
- Cohen, S. *J. biol. Chem.* **257**, 1555–1562 (1982).
- Ross, R. in *Tissue Growth Factors* (ed. Baserga, R.) 133–159 (Springer, Berlin, 1981).
- Rozengurt, E. *Molec. Biol. Med.* **1**, 169–181 (1983).
- James, R. & Bradshaw, R. A. *Rev. Biochem.* (in the press).
- Sporn, M. B. & Todaro, G. J. *New Engl. J. Med.* **303**, 878–880, (1980).
- Waterfield, M. D. *et al. Nature* **304**, 35–39 (1983).
- Doolittle, R. F. *et al. Science* **221**, 275–277 (1983).
- Robbins, K. C., Antoniades, H. N., Devare, S. G., Hunkapiller, M. W. & Aaronson, S. A. *Nature* **305**, 605–608 (1983).
- Deuel, T. F., Huang, J. S., Huang, S. S., Stroobant, P. & Waterfield, M. D. *Science* **221**, 1348–1350 (1983).
- Dulak, N. C. & Temin, H. M. *J. cell. Physiol.* **81**, 161–170 (1973).
- De Larco, J. E. & Todaro, G. J. *Nature* **272**, 356–358 (1978).
- Burk, R. R. *Expl. Cell Res.* **101**, 193–298 (1976).
- Bourne, H. & Rozengurt, E. *Proc. natn. Acad. Sci. U.S.A.* **73**, 4555–4559 (1976).
- De Larco, J. E. & Todaro, G. J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4001–4005 (1978).
- Roberts, A. B., Froluk, C. A., Anzano, M. A. & Sporn, M. B. *Fedn Proc.* **42**, 2621–2626 (1983).
- Ozanne, B., Fulton, R. J. & Kaplan, P. L. *J. cell. Physiol.* **105**, 163–180 (1980).
- Kaplan, P. L., Anderson, M. & Ozanne, B. *Proc. natn. Acad. Sci. U.S.A.* **79**, 485–489 (1982).
- Kaplan, P. L. & Ozanne, B. *Cell* **33**, 931–938 (1983).
- Marquardt, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 4684–4688 (1983).
- Bowen-Pope, D. F., Di Corleto, P. E. & Ross, R. *J. Cell Biol.* **96**, 679–683 (1983).
- Collins, M. K. L., Sinnott-Smith, J. W. & Rozengurt, E. *J. biol. Chem.* **258**, 11689–11693 (1983).
- Marquardt, H. & Todaro, G. J. *J. biol. Chem.* **257**, 5220–5225 (1982).
- Carpenter, G., Stoscheck, C. M., Preston, Y. A. & De Larco, J. E. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5627–5630 (1983).
- Cohen, S., Ushiro, H., Stoscheck, C. & Chinkers, M. *J. biol. Chem.* **257**, 1523–1531 (1982).
- Buhrow, S. A., Cohen, S. & Staros, J. V. *J. biol. Chem.* **257**, 4019–4022 (1982).
- Cohen, S., Fava, R. A. & Sawyer, S. T. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6237–6241 (1982).
- Buhrow, S. A., Cohen, S., Garbers, D. L. & Staros, J. V. *J. biol. Chem.* **258**, 7824–7827 (1983).
- Bishop, J. M., *Rev. Biochem.* **52**, 301–354 (1983).
- Yamamoto, T., Nishida, T., Miyajima, N., Kawai, S., Ooi, T. & Toyoshima, K. *Cell* **35**, 71–78 (1983).
- Adamson, E. D. & Rees, A. R. *Molec. cell. Biochem.* **34**, 129–152 (1981).
- Linsley, P. S., Das, M. & Fox, C. F. in *Membrane Receptors* Vol. B11 (eds Jacobs, S. & Cuatrecasas, P.) 87–113 (Chapman & Hall, London, 1981).
- Schreiber, A. B., Lax, I., Yarden, Y., Eshhar, Z. & Schlessinger, J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7535–7539 (1981).
- Waterfield, M. D. *et al. J. Cell Biochem.* **20**, 149–161 (1982).



35. Kawamoto, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 1337-1341 (1983).
36. Richert, N. D., Willingham, M. C. & Pastan, I. *J. biol. Chem.* **258**, 8902-8907 (1983).
37. Gregoriou, M. & Rees, A. R. *Cell Biol. Int. Rep.* **7**, 539-540 (1983).
38. Schlessinger, J., Lax, I., Yarden, Y., Kanety, H. & Libermann, T. A. in *Receptors and Recognition: Antibodies against Receptors* (ed. Greaves, M. F.) (Chapman and Hall, London, in the press).
39. Fabricant, R. N., De Larco, J. E. & Todaro, G. J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 565-569 (1977).
40. Wrann, M. M. & Fox, C. F. *J. biol. Chem.* **254**, 8083-8086 (1979).
41. O'Keefe, E., Hollenberg, M. D. & Cuatrecasas, P. *Archs biochem. Biophys.* **164**, 518-526 (1974).
42. Gullick, W., Downward, D. J. H., Marsden, J. J. & Waterfield, M. D. *Analyt. Biochem.* (in the press).
43. Ui, N. *Analyt. Biochem.* **97**, 65-71 (1979).
44. Skehel, J. J. & Waterfield, M. D. *Proc. natn. Acad. Sci. U.S.A.* **72**, 93-97 (1975).
45. Waterfield, M. D. & Scrace, G. T. in *Biological/Biomedical Applications of Liquid Chromatography* Vol. 18 (ed. Hawk, G. L.) 135-157 (Dekker, New York, 1981).
46. Bennett, H. P. J., Browne, C. A. & Solomon, S. J. *Liquid Chromat.* **3**, 1353-1365 (1980).
47. Hewick, R. M., Hunkapiller, M. W., Hood, L. E. & Dreyer, W. J. *J. biol. Chem.* **256**, 7990-7997 (1981).
48. Waterfield, M. D., Scrace, G. & Totty, N. in *Practical Protein Biochemistry* (eds Darbre, A. & Waterfield, M. D.) (Wiley, New York, in the press).
49. Wilbur, W. J. & Lipman, D. J. *Proc. natn. Acad. Sci. U.S.A.* **80**, 726-730 (1983).
50. Mayes, E. L. V. & Waterfield, M. D. *EMBO J.* (in the press).
51. Vennström, B. & Bishop, J. M. *Cell* **28**, 135-143 (1983).
52. Kitamura, N., Kitamura, A., Toyoshima, K., Hirayama, Y. & Yoshida, M. *Nature* **297**, 205-208 (1982).
53. Goodfellow, P. N., Banting, G., Waterfield, M. D., Ozanne, B. *Cytogenet. cell. Genet.* **32**, 282 (1982).
54. Shimizu, N., Behzadian, M. A. & Shimizu, Y. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3600-3604 (1980).
55. Kondo, I. & Shimizu, N. *Cytogenet. cell. Genet.* **35**, 9-14 (1983).
56. Spurr, N. *et al. EMBO J.* **3**, 159-163 (1984).
57. Greenberg, M. E. & Edelman, G. M. *Cell* **33**, 767-779 (1983).
58. Groffen, J., Heisterkamp, N., Reynolds, F. H. Jr. & Stephenson, J. R. *Nature* **304**, 167-169 (1983).
59. Hayman, M. J., Ramsay, G. M., Savin, K., Kitchener, G., Graf, T. & Beug, H. *Cell* **32**, 579-588 (1983).
60. Privalsky, M. L., Sealy, L., Bishop, J. M., McGrath, J. P. & Levinson, A. D. *Cell* **32**, 1257-1267 (1983).
61. Beug, H., Palmieri, S., Freudenstein, C., Zentgraf, H. & Graf, T. *Cell* **28**, 907-919 (1982).
62. Graf, T. & Beug, H. *Cell* **34**, 7-9 (1983).
63. Frykberg, L. *et al. Cell* **32**, 227-238 (1983).
64. Fung, Y.-K. T., Lewis, W. G., Crittenden, L. B. & Kung, H.-J. *Cell* **33**, 357-368 (1983).
65. Rabbitts, T. H., Forster, A., Baer, R. & Hamlyn, P. H. *Nature* **306**, 806-809 (1983).
66. Rabbitts, T. H., Hamlyn, P. H. & Baer, R. *Nature* **306**, 760-765 (1983).
67. Gelmann, E. P., Psallidopoulos, M. C., Papas, T. S. & Favera, R. D. *Nature* **306**, 799-803 (1983).
68. Roy-Burman, P., Devi, B. G. & Parker, J. W. *Int. J. Cancer* **32**, 185-191 (1983).
69. Reddy, E. P., Smith, M. J. & Srinivasan, A. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3623-3627 (1983).
70. Bradford, M. M. *Analyt. Biochem.* **72**, 248-254 (1976).
71. Smith, C. M. *et al. Am. J. Obstet. Gynec.* **128**, 190-196 (1977).
72. Laemmli, U. K. *Nature* **227**, 680-685 (1970).
73. Schwartz, D., Tizzard, R. & Gilbert, W. *Cell* **32**, 853-869 (1983).
74. Smart, J. E. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 6013-6017 (1981).
75. Orr, H. T., Lancet, D., Robb, R., Lopez de Castro, J. A. & Strominger, J. L. *Nature* **282**, 266-270 (1979).

## LETTERS TO NATURE

### Interstellar scintillation and ultra-low-frequency gravitational wave observations

J. W. Armstrong

Jet Propulsion Laboratory, Mail Stop 238-737,  
4800 Oak Grove Drive, Pasadena, California 91109, USA

**High-precision pulsar timing measurements have been used<sup>1-4</sup> to set upper limits on the flux of ultra-long period ( $T \geq 3$  years) stochastic gravitational waves. The data used to produce the current limits appear to be limited by timing noise<sup>5,6</sup> that is intrinsic to the pulsars. Recently, however, there have been very low noise timing observations<sup>7</sup> of the 1.5-millisecond pulsar<sup>8</sup>. This suggests that significant sensitivity improvements in searches for ultra-low-frequency (ULF) gravitational waves may be possible. Here the effects of a heretofore secondary noise source—phase scintillations caused by interstellar electron density fluctuations—are calculated and the prospects for improved ULF gravitational wave measurements using pulsar timing are discussed.**

The observed quantity in these experiments is a time series of the difference between the observed and expected pulse arrival times. The time series contains contributions from intrinsic timing noise, gravitational radiation (buffeting of the pulsar and Earth at different epochs), estimation error due to finite signal to noise ratio, and interstellar phase scintillation (ISS). The ISS contribution arises due to dispersion measure fluctuations as irregular clouds of electrons rearrange and drift across the pulsar-Earth line<sup>9,10</sup>. To estimate the contribution to pulsar timing noise measurements a model of the interstellar turbulence spectrum and the relevant propagation theory are required. Observations<sup>10-12</sup> are consistent with an isotropic three-dimensional electron density power spectrum,  $P_{3N}(q)$ , of the form  $P_{3N}(q) = C_N^2 q^{-11/3}$  where  $q$  is the spatial wavenumber. The 'structure constant'  $C_N^2$  is typically  $10^{-5}$  to  $10^{-2} \text{ m}^{-6.67}$  (ref. 11), but can be estimated for the line-of-sight to the 1.5-millisecond pulsar from the measured frequency correlation bandwidth in the dynamic spectrum<sup>8</sup> or the measured temporal broadening of pulses<sup>13</sup>. Using the observed correlation bandwidth ( $\approx 1 \text{ MHz}$  at  $\lambda = 0.21 \text{ m}$  observing wavelength) and the observed pulse broadening ( $73 \mu\text{s}$  at  $\lambda = 0.94 \text{ m}$ ),  $C_N^2$  can be estimated using equations 31 and 41 of ref. 10. For homogeneous turbulence along the pulsar-Earth path of 5 kpc (ref. 14), those

observations yield identical estimates of  $C_N^2 = 8 \times 10^{-5} \text{ m}^{-6.67}$ . Because both decorrelation bandwidth and pulse broadening are diffraction phenomena, those observables are dominated by density fluctuations on scales of the Fresnel zone size ( $\sim 10^{10} \text{ m}$ ). However, direct and indirect measurements of  $P_{3N}$  on the  $10^{11}$ – $10^{13} \text{ m}$  scales relevant to ULF gravitational wave measurements are consistent with an extrapolation of the power at smaller scales; the inferred shape and level of the spectrum on these scales is in fairly good agreement with  $P_{3N}(q) = (10^{-4} \text{ m}^{-6.67}) q^{-11/3}$  (ref. 11).

For gravitational wave searches, noise sources can be characterized either by the power spectrum of fractional frequency fluctuations<sup>2,15</sup> or the Allan variance<sup>16</sup> of the frequency time series<sup>17</sup>. The spectrum of fractional frequency fluctuations caused by ISS,  $S_y(f)$ , can be calculated in three steps. First, ignoring diffraction at the relatively large scales of interest here, the two-dimensional spatial spectrum of phase on the Earth's surface is given by (ref. 9)

$$P_{2\phi}(\kappa) = 2\pi\lambda^2 r_e^2 z C_N^2 \kappa^{-11/3} \quad (1)$$

where  $\kappa$  is the two-dimensional wavenumber,  $\lambda$  is the observing wavelength,  $r_e$  is the classical electron radius, and  $z$  is the pulsar-Earth distance. Second, to convert to a temporal spectrum of phase fluctuations it is assumed that the phase scintillation pattern is 'frozen' in a reference frame that moves with a uniform velocity  $v$  transverse to the pulsar-Earth line. Integrating over the wavenumber component associated with the direction perpendicular to the velocity vector then gives the temporal phase spectrum. Finally, noting that frequency is the temporal derivative of phase, the 'derivative theorem' for Fourier transforms<sup>18</sup> can be applied to obtain  $S_y(f)$ . The result is:

$$S_y(f) = \pi^{-1/6} 2^{-2/3} v^{5/3} \lambda^4 c^{-2} z C_N^2 r_e^2 [\Gamma(4/3)/\Gamma(11/6)] f^{-2/3} \quad (2)$$

where  $c$  is the speed of light,  $v$  is the velocity of the pattern,  $\Gamma$  is the gamma function and  $f$  is the temporal frequency.

Equation (2) can be used to predict ISS noise for ULF gravitational wave observations. Figure 1 shows the current observational limits to a ULF gravitational wave background<sup>2-4</sup>, the envelope of stochastic gravitational wave spectra having closure density<sup>19</sup>, and plots of equation (2) evaluated with the following parameters:  $\lambda = 0.21 \text{ m}$  and  $0.06 \text{ m}$ ,  $z = 5 \text{ kpc}$ ,  $C_N^2 = 8 \times 10^{-5} \text{ m}^{-6.67}$ ,  $v = 100 \text{ km sec}^{-1}$ . Also shown in Fig. 1 is the  $S_y(f)$  that would be observed for daily observations of a pulsar having white intrinsic timing noise with spectral density  $9 \times 10^4 \mu\text{s}^2 \text{ Hz}^{-1}$ . That is, the normalization is set by  $1 \mu\text{s}$  (r.m.s.) white timing noise distributed over the Nyquist band<sup>18</sup> 0–0.5 cycles  $\text{d}^{-1}$  set by the daily sampling interval.