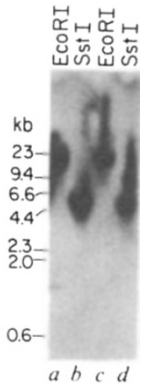

Fig. 4 Southern blot of rat genomic DNA probed with the TGF-α cDNA clone. DNA was isolated from adult female rat kidney (lanes *a*, *b*) and FeSV-transformed Fisher rat embryo cells (lanes *c*, *d*) and restricted with the indicated enzymes. Digested DNA (20 μg) was electrophoresed on a 1% agarose gel, transferred to nitrocellulose by the method of Southern[23] and probed with the nick-translated *Eco*RI insert from prTGF$_{0,2}$. Hybridization and washes were carried out as described in Fig. 3 legend. The markers (in kb) are *Hin*dIII-digested λ DNA.

necessary to code for a 50-amino acid polypeptide. Northern blots of poly(A)$^+$ RNA obtained from FeSV-transformed rat cells, when probed with the *Eco*RI insert from prTGF$_{0,2}$, yielded a single band of ~4.5 kb (Fig. 3). In addition, as sequence analysis of the 3′ end of the prTGF$_{0,2}$ insert revealed a 20-residue poly(A) tract, the cloned cDNA apparently corresponds to the second half of the TGF-α mRNA. The 50-amino acid TGF-α is encoded by a sequence located roughly 2.5 kb from the 5′ end of the mRNA.

Northern blot analysis of poly(A)$^+$ RNA from normal rat brain, when probed with the TGF-α clone, revealed a single band identical in size to that obtained from the FeSV-transformed cells (Fig. 3). The level of TGF-α-specific mRNA relative to total poly(A)$^+$ RNA in brain was roughly 10–20-fold lower than that in FeSV-transformed cells (data not shown). In addition, we have observed TGF-α mRNA in rat kidney and liver in amounts similar to that in brain (data not shown). The presence of TGF-α-specific mRNA in tissues raises interesting questions regarding the role of this polypeptide in normal cells.

As described above, TGF-α is structurally homologous to EGF, and both bind, in an equivalent fashion, the EGF receptor. This similarity is extended further by a comparison of the cDNAs encoding EGF[9,10] and TGF-α which indicates that both polypeptides are synthesized as larger proteins encoded by mRNAs of 4.5–5.0 kb. In addition, both growth factors share striking homology with a 140-base pair vaccinia virus protein of unknown function (J. Brown *et al.*, accompanying paper[24]). As both growth factors appear, therefore, to belong to a family of EGF-like polypeptides, we used the TGF-α cDNA to screen Southern blots of rat genomic DNA to identify homologous genes in this family. In moderate stringency conditions, however, only one (*Sst*I) or two (*Eco*RI; the second band of ~1.0 kb may not be visible) bands appeared with each restriction enzyme tested (Fig. 4), indicating that the TGF-α gene is probably unique in rats and is not highly homologous to genes encoding other related polypeptides. In addition, a comparison of the results obtained with DNAs isolated from kidney and FeSV-transformed cells indicated that elevated levels of TGF-α mRNA in the latter are not the result of gross amplification or rearrangement of the gene.

We thank Jane Ranchalis for providing FeSV-transformed Fisher rat embryo cells, Daniel Twardzik, Hans Marquardt, Peter Linsley, Debra Bryant, Hans Neurath and Kenneth Walsh for helpful discussions, and Lynda Taylor for assistance in preparing the manuscript.

Note added in proof: After this manuscript had been submitted, the cloning of human TGF-α was reported[25].

Received 1 August; accepted 7 November 1984.

1. Todaro, G. J., DeLarco, J. E., Fryling, C., Johnson, P. A. & Sporn, M. B. *J. supramolec. Struct.* **15**, 287–301 (1981).
2. DeLarco, J. E. & Todaro, G. J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4001–4005 (1978).
3. Pike, L. J. *et al. J. biol. Chem.* **257**, 14628–14631 (1983).
4. Marquardt, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 4684–4688 (1983).
5. Savage, C. R. Jr, Inagami, T. & Cohen, S. *J. biol. Chem.* **247**, 7612–7621 (1972).
6. Gregory, H. *Nature* **257**, 325–327 (1975).
7. Marquardt, H., Hunkapiller, H. W., Hood, L. E. & Todaro, G. J. *Science* **223**, 1079–1082 (1984).
8. Anzano, M. A. *et al. Cancer Res.* **42**, 4776–4778 (1982).
9. Gray, A., Dull, T. J. & Ullrich, A. *Nature* **303**, 722–725 (1983).
10. Scott, J. *et al. Science* **221**, 236–240 (1983).
11. Twardzik, D. R., Todaro, G. J., Reynolds, F. H. Jr & Stephenson, J. R. *Virology* **124**, 201–207 (1983).
12. Huynh, T., Young, R. & Davis, R. in *Practical Approaches in Biochemistry* (ed. Glover, D.) (IRL, Oxford, 1984).
13. Dente, L., Cesareni, G. & Cortese, R. *Nucleic Acids Res.* **11**, 1645–1655 (1983).
14. Linsley, P. S., Hargreaves, W. R., Twardzik, D. R. & Todaro, G. J. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
15. Naughton, M. A. & Sanger, F. *Biochem. J.* **78**, 156–162 (1961).
16. Matsubara, H. *Meth. Enzym.* **19**, 642–651 (1970).
17. Kreil, G. *A. Rev. Biochem.* **50**, 317–348 (1981).
18. Okayama, H. & Berg, P. *Molec. cell. Biol.* **2**, 161–167 (1982).
19. Smith, A. J. H. *Meth. Enzym.* **65**, 560–580 (1980).
20. Glisin, V., Crkvenjakov, R. & Byns, C. *Biochemistry* **13**, 2633–2638 (1974).
21. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294–5299 (1979).
22. Maniatis, T., Fritsch, E. F. & Sambrook, J. in *Molecular Cloning* (Cold Spring Harbor Laboratory, New York, 1982).
23. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
24. Brown, J. P., Twardzik, D. R., Marquardt, H. & Todaro, G. J. *Nature* **313**, 491–492 (1985).
25. Derynck, R., Roberts, A. B., Winkler, M. E., Chen, E. Y. & Goeddel, D. V. *Cell* **38**, 287–297 (1984).

Vaccinia virus encodes a polypeptide homologous to epidermal growth factor and transforming growth factor

Joseph P. Brown*†, Daniel R. Twardzik*, Hans Marquardt* & George J. Todaro*†

* Oncogen, 3005 First Avenue, Seattle, Washington 98121, USA
† Department of Pathology, University of Washington Medical School, Seattle, Washington 98195, USA

Epidermal growth factor (EGF) and transforming growth factor type I (TGF) are polypeptides of 53 and 50 amino acid residues, respectively[1,2]. Both bind to EGF receptor, a 1,200-residue transmembranous glycoprotein[3], leading to phosphorylation of the receptor, enhancement of its tyrosine-specific kinase activity and ultimately to stimulation of cell growth[4,5]. We report here that a 140-residue polypeptide encoded by one of the early genes of vaccinia virus (VV)[6] is related closely to EGF and TGF. The presence of putative signal and transmembranous sequences further suggests that the viral protein might be an integral membrane protein, but that, as in the case of EGF itself[7,8], the membrane-associated form may be the precursor of a soluble growth factor. Production of EGF-like growth factors by virally infected cells could account for the proliferative diseases associated with members of the poxvirus family such as Shope fibroma virus[9], Yaba tumour virus[10], and molluscum contagiosum virus (MCV)[11].

The Dayhoff protein sequence library of 2,676 sequences (Protein Identification Resource, Georgetown, Washington, DC) was searched for sequences related to rat TGF (rTGF). The three most closely related sequences were mouse EGF (mEGF)[7,8] and human EGF (hEGF, urogastrone)[12], which are known to be homologous to TGF[13], and residues 45–85 of a 140-residue polypeptide encoded by VV[6] (Fig. 1). Fifteen residues of the VV polypeptide match residues in rTGF and, after insertion of a single gap, the viral polypeptide shares 19 residues with both mEGF and hEGF; a gap must be inserted in the same position in TGF to align it with EGF. The viral protein is thus as similar to EGF as is rTGF, which has 16 residues in common with mEGF and 20 with hEGF. The statistical significance of the homology can be determined as follows. Of the 16 residues of mEGF which are conserved in rTGF, 13

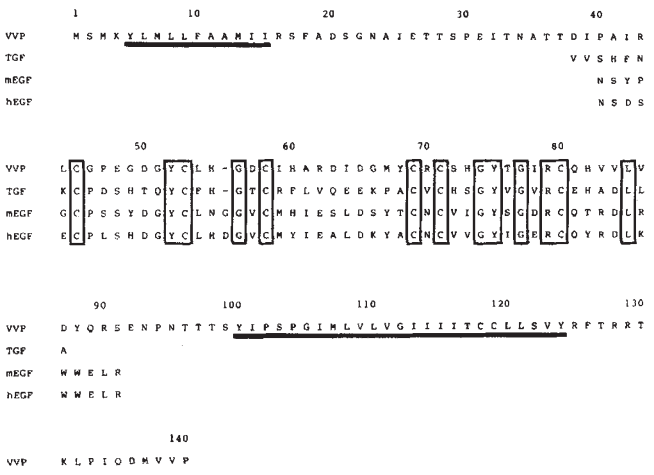

Fig. 1 Alignment of the VV polypeptide (VVP), rTGF, mEGF and hEGF. The sequences are shown in their entirety, numbered from the N-terminus of VVP. Residues conserved in all four sequences are boxed. Putative signal and transmembranous sequences are underlined. A, Ala; R, Arg; N, Asn; D, Asp; C, Cys; Q, Gln; E, Glu; G, Gly; H, His; I, Ile; L, Leu; K, Lys; M, Met; F, Phe; P, Pro; S, Ser; T, Thr; W, Trp; Y, Tyr; V, Val.

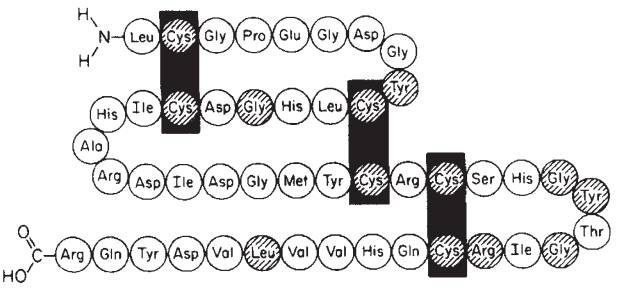

Fig. 2 Proposed structure of the VV polypeptide that would be produced by cleavage of the precursor at Arg 43 and Arg 90. Disulphide bonds (boxed) are based on those of EGF. Residues that are identical in the VV polypeptide, mEGF, hEGF and rTGF are cross-hatched.

are conserved in the VV polypeptide, whereas of the remaining 37 residues of mEGF, only six are conserved in the VV polypeptide. Fisher's exact test shows that the probability of this being due to chance is less than 0.0001. In addition to EGF, the EGF precursor encodes eight sequences that resemble EGF in their pattern of cysteine residues and their predicted folding patterns[7,8,14]; EGF-related sequences are also present in the low-density lipoprotein receptor[15] and the blood coagulation zymogens factor IX, factor X and protein C (ref. 14). It thus appears that these sequences and EGF diverged from a common ancestor[14]. However, at least five gaps must be inserted to align the VV polypeptide with any of these sequences, and the VV polypeptide is most closely related to EGF and TGF.

The EGF-homologous region of the VV polypeptide differs from the rest of the molecule in its amino acid composition, particularly in that it contains 6 of the 8 cysteine residues and 7 of the 10 glycine residues. These residues may have a structural role. The six cysteine residues correspond to the six cysteine residues that, in EGF, form three disulphide bridges, resulting in three tight loops in the polypeptide chain[16]. Three of the glycine residues are conserved in EGF and TGF, and together with the four aspartic acid residues and the single proline residue present in this region of the VV polypeptide, they may allow the chain to fold back on itself within the right constraints imposed by the disulphide bridges[14,16,17]. Two regions of particularly high homology are evident; these are residues 50–56, all of which are identical to residues in the corresponding positions in either EGF or TGF, and residues 73–82, seven of which are conserved in EGF or TGF. We therefore conclude that the VV polypeptide probably resembles EGF and TGF not only in its amino acid sequence, but also in its pattern of disulphide bridges and thus in its three-dimensional structure.

Further analysis of the sequence of the viral protein revealed the presence of stretches of uncharged and hydrophobic residues near the N-terminus between residues 5 and 15 and near the C-terminus between residues 100 and 124. This arrangement of hydrophobic regions is similar to that seen in many integral membrane glycoproteins, which have an N-terminal hydrophobic signal sequence and a C-terminal transmembranous sequence. The transmembranous sequence is followed commonly by several basic residues and in the VV polypeptide, residues 125, 128 and 129 are arginine. The EGF precursor itself has signal and transmembranous sequences; it has been suggested that it exists on the surface of the cells of its origin anchored by the hydrophobic segment, EGF being released by proteolytic cleavage[14]. The transmembranous peptides of both the EGF precursor and the VV polypeptide are located ~20 residues from the end of the EGF-like sequence; cleavage of the viral polypeptide at Arg 43 and Arg 90 would lead to release of a soluble polypeptide of 47 residues (Fig. 2).

As the VV polypeptide is structurally as similar to EGF as is TGF, we believe that, like TGF, it might be able to bind the EGF receptor and act as a growth factor. This suggests several possible roles for the VV polypeptide. If it is located on the surface of the virion, it might mediate binding of the virus to cells expressing the EGF receptor and also stimulate the host cell. If, on the other hand, a soluble EGF-like peptide is released from VV-infected cells, neighbouring cells might be stimulated. It is intriguing to speculate that involvement of the EGF receptor in VV infection might contribute towards the propensity of the virus to infect epidermal cells, while production of EGF-like growth factors might account for the cellular proliferation associated with poxviruses both *in vivo*[9–11] and *in vitro*[18–20].

In conclusion, the structural homology between the VV polypeptide and EGF and TGF indicates clearly an evolutionary relationship. Whether a functional relationship exists remains to be confirmed experimentally. Nevertheless, our observation, together with the many reports of the capture of cellular genes by retroviruses, including genes for platelet-derived growth factor[21] and EGF receptor[22], suggest that viral capture of cellular genes coding for growth factors and receptors might be more common than is realized generally and might not be restricted to the retroviruses.

Note added in proof: Blomquist *et al.*[23] have recently reported findings similar to those described here.

Received 1 August; accepted 13 November 1984.

1. Carpenter, G. & Cohen, S. *A. Rev. Biochem.* **48**, 193-216 (1979).
2. Marquardt, H., Hunkapiller, M. W., Hood, L. E. & Todaro, G. J. *Science* **223**, 1079-1082 (1984).
3. Ullrich, A. *et al. Nature* **309**, 418-425 (1984).
4. Cohen, S. in *Biological Response Mediators and Modulators* (ed. August, J. T.) 7-12 (Academic, New York, 1983).
5. Tam, J. P., Marquardt, H., Rosberger, D. F., Wong, T. W. & Todaro, G. J. *Nature* **309**, 376-378 (1984).
6. Vankatesan, S., Gershowitz, A. & Moss, B. *J. Virol.* **44**, 637-646 (1982).
7. Scott, J. *et al. Science* **221**, 236-240 (1983).
8. Gray, A., Dull, T. J. & Ullrich, A. *Nature* **303**, 722-725 (1983).
9. Shope, R. E. *J. exp. Med.* **56**, 793-822 (1932).
10. Niven, J. S. F., Armstrong, J. A., Andrewes, C. H., Pereira, H. G. & Valentine, R. C. *J. Path. Bact.* **81**, 1-14 (1961).
11. Postlethwaite, R. *Archs envir. Hlth* **21**, 432-452 (1970).
12. Gregory, H. & Preston, B. M. *Int. J. Peptide Protein Res.* **9**, 107-118 (1977).
13. Marquardt, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 4684-4688 (1983).
14. Doolittle, R. F., Feng. D. F. & Johnson, M. *Nature* **307**, 558-560 (1984).
15. Russell, D. W. *et al. Cell* **37**, 577-585 (1984).
16. Savage, C. R., Hash, J. H. & Cohen, S. *J. biol. Chem.* **248**, 7669-7692 (1973).
17. Chou, P. Y. & Fasman, G. D. *A. Rev. Biochem.* **47**, 251-276 (1978).
18. Rouhandeh, H. & Vafai, A. *Virology* **120**, 71-92 (1982).
19. Koziorowska, J., Wlodarski, K. & Mazurowa, N. *J. natn. Cancer Inst.* **46**, 225-241 (1971).
20. Barbanti-Brodano, G., Mannini-Palenzona, A., Varoli, O., Portolani, M. & La Placa, M. *J. gen. Virol.* **24**, 237-246 (1974).
21. Waterfield, M. D. *et al. Nature* **304**, 35-39 (1983).
22. Downward, J. *et al. Nature* **307**, 521-527 (1984).
23. Blomquist, M. C., Hunt, L. T. & Barker, W. C. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7363-7367 (1984).