had the additional scans, there were still significant increases in temporopolar blood during the production of anticipatory anxiety.
20. B. R. Kaada, *Handbook of Physiology*, Section 1, *Neurophysiology*, J. Field and H. W. Magoun, Eds. (American Physiological Society, Washington, DC, 1960), vol. 2, pp. 1345–1372.
21. W. Penfield and H. Jasper, *Epilepsy and the Functional Anatomy of the Human Brain* (Little, Brown, Boston, 1954).
22. A. Kling and H. D. Steklin, *Brain Behav. Evol.* **13**, 216 (1976).
23. E. Halgren and R. Walter, *Brain* **101**, 83 (1978); P. Gloor, A. Olivier, L. F. Quesney, F. Andermann, S. Horowitz, *Ann. Neurol.* **12**, 129 (1982); E. Strauss, A. Risser, M. W. Jones, *Arch. Neurol.* **39**, 626 (1982).
24. M. Wall *et al.*, *J. Clin. Psychiatry* **46**, 143 (1985); M. Wall, D. Mielke, J. S. Luther, *ibid.* **47**, 219 (1986); A. M. Ghadirian *et al.*, *ibid.*, p. 270.
25. M. M. Mesulam and E. J. Mufson, *J. Comp. Neurol.* **212**, 1 (1982); M. A. Moran, E. J. Mufson, M. M. Mesulam, *ibid.*, **256**, 88 (1987).
26. There was no significant difference between the left-to-right ratio of parahippocampal blood in the second and third scans ($P = 0.97$). Thus, the abnormal parahippocampal asymmetry in patients with panic disorder seems to be unrelated to a state of anticipatory anxiety.
27. J. Talairach and G. Szikla, *Atlas d'Anatomie Stereotaxique du Telencephale* (Masson, Paris, 1967).
28. We thank J. Stern and J. Price for helpful suggestions, L. Lich for technical support, and S. Sirkin and J. Woodward for secretarial assistance. Supported by National Institute of Mental Health Physician Scientist Award MH-00615 (E.M.R.); National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) Teacher Investigator Award NS-00904 (P.T.F.); NINCDS grant NS-06833; National Heart, Lung, and Blood Institute grant HL-13851; and a grant from the Upjohn company.

29 August 1988; accepted 11 November 1988

Structure and Function of Human Amphiregulin: A Member of the Epidermal Growth Factor Family

MOHAMMED SHOYAB, GREGORY D. PLOWMAN, VICKI L. MCDONALD, J. GARRETT BRADLEY, GEORGE J. TODARO

The complete amino acid sequence of amphiregulin, a bifunctional cell growth modulator, was determined. The truncated form contains 78 amino acids, whereas a larger form of amphiregulin contains six additional amino acids at the amino-terminal end. The amino-terminal half of amphiregulin is extremely hydrophilic and contains unusually high numbers of lysine, arginine, and asparagine residues. The carboxyl-terminal half of amphiregulin (residues 46 to 84) exhibits striking homology to the epidermal growth factor (EGF) family of proteins. Amphiregulin binds to the EGF receptor but not as well as EGF does. Amphiregulin fully supplants the requirement for EGF or transforming growth factor–α in murine keratinocyte growth, but it is a much weaker growth stimulator in other cell systems.

THE LIST OF PEPTIDE GROWTH REGulators has been expanding rapidly. These factors participate in various physiological and pathological conditions, such as cellular communication, growth and development, embryogenesis, immune response, hematopoiesis, cell survival and differentiation, inflammation, tissue repair and remodeling, atherosclerosis, and cancer (*1*). The isolation, characterization, and mechanism of action of regulatory factors for growth and differentiation are of current interest because of the potential use of such regulatory factors in the diagnosis, prognosis, and therapy of neoplasia and because of what these factors reveal about the basic mechanism of normal cellular proliferation and the unrestrained growth of cancer cells. We have recently reported the isolation of a novel glycoprotein termed amphiregulin (AR), which inhibits growth of A431 human epidermoid carcinoma and other human tumor cells and stimulates proliferation of human fibroblasts and other normal and tumor cells (*2*). AR was isolated from serum-free conditioned medium of MCF-7 human breast carcinoma cells that had been treated with 12-*O*-tetradecanoylphorbol-13-acetate (*2*). We now report the complete amino acid sequence of amphiregulin and compare its biological properties with those of the other members of the epidermal growth factor (EGF) family proteins.

AR was purified to homogeneity as described (*2*). The homogeneous AR was used for all the chemical and biological studies reported here. The amino acid sequence of human AR (Fig. 1) was determined by automated Edman degradation of *N*-glycanase–treated, reduced, and *S*-pyridylethylated AR (NG-SPE-AR) and of peptide fragments obtained by cleavage of NG-SPE-AR with various endopeptidases. The carboxyl-terminal analysis of NG-SPE-AR was performed with carboxypeptidase P (*Penicillium janthinellum*). The amino-terminal analysis of NG-SPE-AR revealed the presence of two sequences, one starting at residue 1, serine, and the other starting at residue 7, valine (Fig. 1). The yield of the larger form of AR was about 20% of that of the truncated form. The larger AR thus contains six additional amino acids at the amino terminal of the truncated form of AR. The larger form of AR and the truncated AR are single chain polypeptides of 84 and 78 residues, with a calculated molecular weight of 9759 and 9060, respectively (Fig. 1). Both forms of AR have a similar carboxyl-terminal sequence as determined by carboxypeptidase P cleavage (Fig. 1), and both are biologically active.

The sequence of AR was compared with all proteins in the National Biomedical Research Foundation database (release 15, containing 6796 protein sequences), Genetic Sequence Data Bank (Bolt Beranek and Newman, Los Alamos National Laboratory; release 54) and the European Molecular Biology Laboratory DNA sequence library (release 13). These computer-aided searches revealed that AR is a novel protein and a member of the EGF family. This family includes EGF (mouse, human, and rat) (*3–5*), transforming growth factor–α (TGF-α) (*6, 7*), and poxvirus growth factors [vaccinia (VGF), myxoma (MGF), and Shope fibroma (SFGF)] (*8–10*). Tissue-type plasminogen activator (*11*), the mammalian clotting factors IX and X (*12*), the low-density lipoprotein receptor (*13*), bovine protein C (*14*), human proteoglycan core protein (*15*), product of *Drosophila notch* gene (*16*), product of *lin 12* gene (*17*), the product of cell lineage–specific gene of sea urchin *Strongylocentrotus purpuratus* (*18*), cytotactin (*19*), and product of *Pfs* gene of *Plasmodium falciparum* (*20*) also contain EGF-like domains. Alignment of AR structure with the structure of EGF-like growth factors and with other members of EGF-like proteins (Fig. 2) reveals that AR, like other members of the family, contains the hallmark six essential cysteine residues, maintains conservation of cysteine residue spacing in the pattern $CX_7CX_4CX_{10}CX_1CX_8C$, and also contains some of the characteristic and conserved amino acids. AR falls between the members of the growth factor family that look like EGF and TGF-α and those that look like the poxvirus-encoded growth factors (MGF and SFGF), especially in the use of asparagine. The amino-terminal sequence of AR has some analogy with the amino-terminal sequences of the TGF-α's (*6, 7*), VGF (*8*), and MGF (*9*) in that it is rich in prolines, serines, and threonines and, like TGF-α and VGF, has potential N-linked glycosylation sites as well as the possibility for O-linked glycosylation in the region rich in serines, threonines, and prolines. Unlike MGF and SFGF, AR does not have any potential glycosylation site within the growth factor domain of the molecule. On the basis of homology with mouse EGF (*3*) and perfect alignment of six cysteine residues, one would expect the pres-

Oncogen, 3005 First Avenue, Seattle, WA 98121.

ence of three intrachain disulfide bonds in AR involving cysteine residues 46 and 59, 54 and 70, and 72 and 81.

AR is an extremely hydrophilic protein, especially the amino-terminal half of the molecule up to residue 45. A 23–amino acid stretch from residue 23 through 45 contains only five different amino acids (ten lysines, four arginines, four asparagines, three prolines, and two glycines). A tetrapeptide Arg-Lys-Lys-Lys is repeated twice (residues 26 to 29 and 40 to 43) in AR. Such sequences have been reported to serve as a nucleus targeting signal (*21*). The hydropathy profile of AR exhibits little similarity with those of the other members of the EGF family.

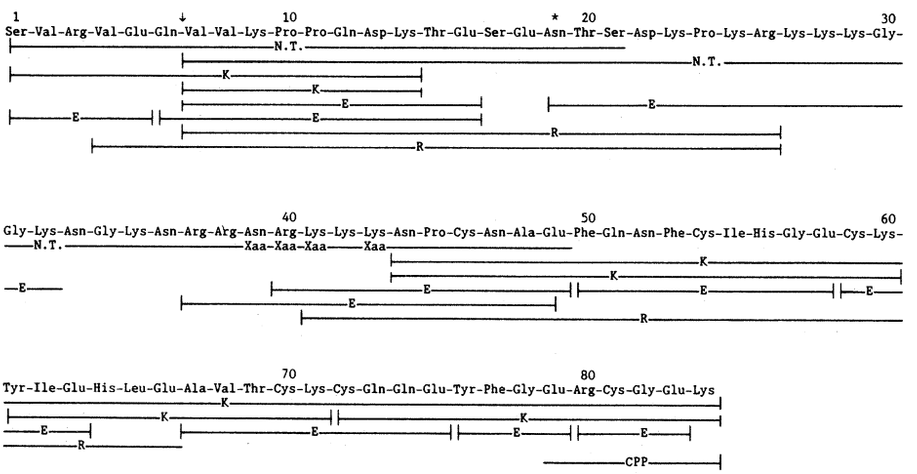


Fig. 1. Amino acid sequence of AR and schematic outline of the data supporting the sequence. The sequence of unfragmented NG-SPE-AR is denoted by N.T. Peptides obtained by cleavage with endopeptidase-Lys-C (K), with endopeptidase-Arg (R), and with endopeptidase-Glu, *Staphylococcus aureus* V8 protease (E) are indicated. CPP denotes the carboxyl-terminal sequence determined by digestion of AR with carboxypeptidase P. Residues identified with Edman degradation or by amino acid analysis are indicated by lines. Vertical bars show beginnings and endings of fragments. Lines without two vertical bars indicate incomplete sequences; ↓ indicates the start of truncated AR; and * indicates potential glycosylation site. AR was reduced with 2-mercaptoethanol and alkylated with 4-vinylpyridine. SPE-AR was purified by reversed phase high-performance liquid chromatography (rpHPLC). SPE-AR was treated with *N*-glycanase to remove N-linked oligosaccharides, and NG-SPE-AR was purified by rpHPLC. NG-SPE-AR was cleaved with various endopeptidases, and the resulting peptides were separated by an rpHPLC C_8 column. Peptide sequences were determined with an Applied Biosystems model 475A gas-phase sequencer. Identification of phenylthiohydantoin amino acid derivatives was carried out, on line, on a model 120A analyzer (Applied Biosystems). For carboxyl-terminal analysis, NG-SPE-AR was incubated with CPP, portions were withdrawn at various times, and the reaction was terminated. Released amino acids were derivatized with phenyl isothiocyanate and phenylthiocarbamyl amino acid derivates were analyzed and quantitated by using micro amino acid derivatizer and analyzer (Applied Biosystem, model number 420-A0-03).

The binding properties of AR were compared with those of mouse EGF in radioreceptor assays (Fig. 3A). AR inhibited the binding of ^{125}I-labeled EGF to A431 cells as well as to A431 plasma membranes. A 50% inhibition of ^{125}I-EGF binding to fixed cells and membranes was seen at about 1.1 and 1.8 n*M* EGF, respectively, whereas a 50% reduction in EGF binding to cells and membranes was seen at approximately 1.8 and 5.7 n*M* AR, respectively. Unlabeled EGF completely inhibited the ^{125}I-EGF–receptor interaction at higher concentrations in both systems (Fig. 3A). However, the maximum competition with AR was 75% and 50% for binding to cells and membranes, respectively (Fig. 3A). The competition curves for AR were not parallel to that seen with EGF. These results suggest that AR has a lower affinity for EGF receptors on A431 cells than does EGF itself. Structural differences between AR and EGF might explain the binding data shown in Fig. 3. It is also possible that AR might have its specific receptor closely related to the EGF receptor.

EGF or TGF-α induce anchorage-independent growth of rat kidney cells NRK-SA6 in the presence of TGF-β (*22*). EGF induced anchorage-independent growth of NRK cells in a dose-dependent manner in the presence of TGF-β, whereas AR was found to be a noninducer of colony formation in soft agar of NRK cells (Fig. 3B). The continued growth of a murine keratinocyte cell line, Balb/MK, is dependent on EGF or TGF-α (*23*). Balb/MK cells did not proliferate in the absence of AR or EGF. However, these cells proliferated equally well in the presence of AR or of EGF (Fig. 3C). Thus,

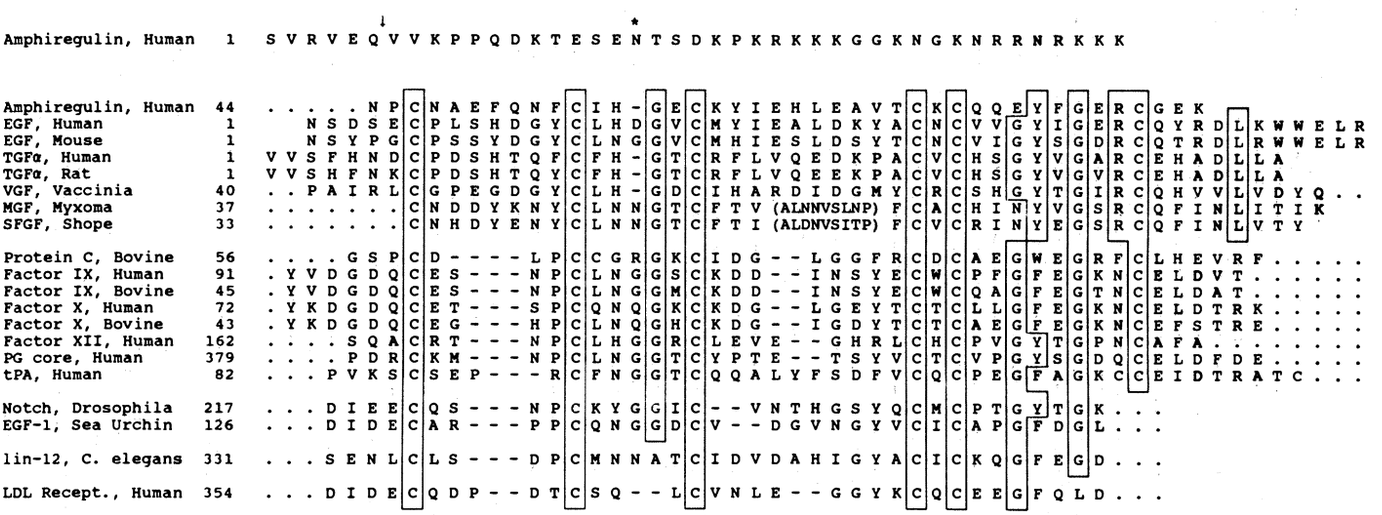
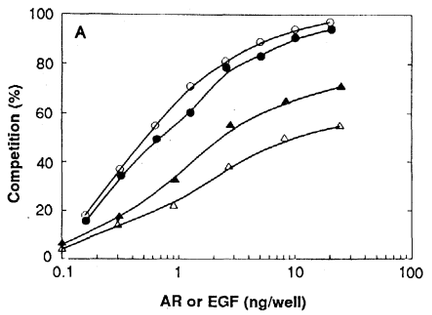
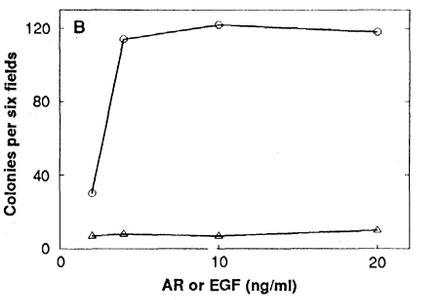


Fig. 2. Alignment of AR sequence with other EGF-like proteins. Amino acids are represented by standard one-letter symbols (*24*). Only residues appearing in eight or more proteins are boxed. Hyphens indicate gap introduced to maximize homology. Dots at the beginning and the end of sequences indicate the use of only partial sequence of a given protein. Numbers at the beginning of every sequence indicate the number of amino acid residues within the total protein sequence; ↓, beginning of the sequence of the truncated form of AR; and *, potential glycosylation site.





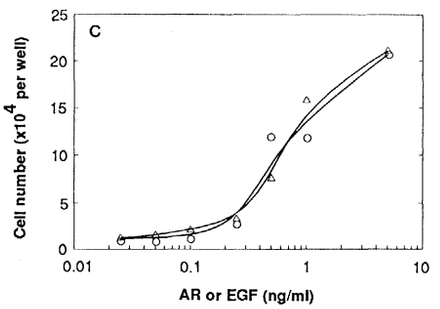


Fig. 3. (**A**) Competition of ^{125}I-EGF binding to the fixed A431 cells (solid symbols) or A431 plasma membranes (open symbols) by murine EGF and AR. EGF was radioiodinated with ^{125}I as described (*25*). The binding assays were performed either in 48-well tissue culture plates when Formalin-fixed A431 cells were used as described (*26*) or by immobilizing plasma membranes onto 96-well poly(vinyl chloride) plates as described (*27*). The binding assays used 4 ng of ^{125}I-labeled mouse EGF per milliliter, containing ~1.9×10^5 dpm. Samples of 100 and 50 μl were used per well for assays with fixed cells and membranes, respectively. Circles indicate EGF, and triangles indicate AR. (**B**) Effect of EGF and AR on NRK-SA6 cell colony formation in soft agar in the presence of TGF-β (1 ng/ml). A 0.38-ml base layer of 0.5% agar (Agar Noble, Difco Laboratories, Detroit) in Dulbecco's minimum essential medium containing 10% heat-inactivated fetal bovine serum (FBS) was added to 24-well Costar tissue culture plates. A 0.3% agar (0.38 ml) containing the same medium-FBS mixture, 6×10^3 to 12×10^3 test cells, and the factors to be tested were overlaid on the basal layer of agar. The plates were incubated at 37°C in the humidified atmosphere of 5% CO_2 in air. Colonies were enumerated unfixed and unstained, and the number of colonies was scored between days 7 and 10. Colonies were defined as a cluster of at least eight cells. Circles, EGF; and triangles, AR. (**C**) Effect of AR and EGF on the growth of murine keratinocytes. Balb/MK cells were plated at 1×10^4 cells per well in 1 ml of low calcium medium (*23*) in 24-well Costar plates (area ~2 cm^2 per well) and incubated overnight at 37°C. Then media were removed and replaced with 1 ml of medium containing various concentrations of AR or EGF in triplicate. The control wells received only medium without any AR or EGF. Plates were incubated at 37°C for 4 days, then medium was removed, wells were rinsed two times with 1 ml of phosphate-buffered saline, and the cells were detached with trypsin-EDTA and counted. Circles, EGF; and triangles, AR.

AR can supplant the EGF requirement in these cells. These results indicate that, like EGF and TGF-α, AR acts as a growth stimulator, but is much weaker on some cells (normal rat kidney) and comparable on others (murine keratinocytes).

Available structural data should allow studies on the cloning, structure, topology, expression, and regulation of amphiregulin gene in both the physiological and pathological conditions. These studies may also provide clues to design agonists and antagonists of this bifunctional growth regulator.

REFERENCES AND NOTES

1. D. Evered, J. Nugent, J. Whelan, Eds., *Ciba Found. Symp.* **116** (1985); G. Carpenter and S. Cohen, *Annu. Rev. Biochem.* **48**, 193 (1979); E. Rozengurt, *Science* **234**, 161 (1986); A. B. Pardee, *Cancer Res.* **47**, 1488 (1987); A. S. Goustin, E. B. Leof, G. D. Shipley, H. L. Moses, *ibid.* **46**, 1015 (1986); M. B. Sporn and A. B. Roberts, *J. Clin. Invest.* **78**, 329 (1986); L. Sachs, *Sci. Am.* **254**, 40 (January 1986); L. J. Old, *ibid.* **258**, 59 (May 1988); B. Beutler and A. Cerami, *N. Engl. J. Med.* **316**, 379 (1987); J. E. Groopman, *Cell* **50**, 5 (1987); M. J. Clemens and M. A. McNurlan, *Biochem. J.* **226**, 345 (1985).
2. M. Shoyab, V. L. McDonald, J. G. Bradley, G. J. Todaro, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6528 (1988).
3. C. R. Savage, Jr., T. Inagami, S. Cohen, *J. Biol. Chem.* **247**, 7612 (1972); C. R. Savage, Jr., J. H. Hash, S. Cohen, *ibid.* **248**, 7669 (1973); R. F. Doolittle, D. F. Feng, M. S. Johnson, *Nature* **307**, 558 (1984).
4. H. Gregory, *Nature* **257**, 325 (1975).
5. R. J. Simpson *et al.*, *Eur. J. Biochem.* **153**, 629 (1985).
6. G. Todaro, C. Fryling, J. DeLarco, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5258 (1980); H. Marquardt, M. W. Hunkapiller, L. E. Hood, G. J. Todaro, *Science* **223**, 1079 (1984).
7. R. Derynck *et al.*, *Cell* **38**, 287 (1984); D. Lee, T. Rose, N. Webb, G. J. Todaro, *Nature* **313**, 489 (1985).
8. M. C. Blomquist, L. C. Hunt, W. C. Barker, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7363 (1984); J. P. Brown, D. R. Twardzik, H. Marquardt, G. J. Todaro, *Nature* **313**, 491 (1985).
9. C. D. Porter and L. C. Archard, *J. Gen. Virol.* **68**, 673 (1987).
10. W. Chang, C. Upton, S.-L. Hu, A. F. Purchio, G. McFadden, *Mol. Cell. Biol.* **7**, 535 (1987).
11. D. Pennica *et al.*, *Nature* **301**, 214 (1983).
12. K. Kurachi and E. W. Davie, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6461 (1982); K. Katayama *et al.*, *ibid.* 76, 4990 (1979); B. A. McMullen *et al.*, *Biochemistry* **22**, 2875 (1983); D. L. Enfield *et al.*, *ibid.* **19**, 659 (1980); K. Titani *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 3082 (1975).
13. D. W. Russell *et al.*, *Cell* **37**, 577 (1984).
14. P. Fernlund and J. Stenflo, *J. Biol. Chem.* **257**, 12170 (1982).
15. T. Krusius, K. R. Gehlsen, E. Ruoslahti, *ibid.* **262**, 13120 (1987).
16. K. A. Wharton, K. M. Johansen, T. Xu, S. Artavanis-Tsakonas, *Cell* **43**, 567 (1985); S. Kidd, M. R. Kelley, M. W. Young, *Mol. Cell Biol.* **6**, 3095 (1986).
17. I. Greenwald, *Cell* **43**, 583 (1985).
18. D. A. Hursh, M. E. Andrews, R. A. Raff, *Science* **237**, 1487 (1987).
19. F. S. Jones *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2186 (1988).
20. D. C. Kaslow *et al.*, *Nature* **333**, 74 (1988).
21. Y. Yoneda *et al.*, *Science* **242**, 275 (1988).
22. A. B. Robert *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 5339 (1981).
23. B. E. Weissman and S. A. Aaronson, *Cell* **32**, 599 (1983); R. J. Coffey *et al.*, *Cancer Res.* **48**, 1596 (1988).
24. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
25. K. Barridge, *Methods Enzymol.* **50**, 54 (1978).
26. G. Carpenter, L. King, Jr., S. Cohen, *J. Biol. Chem.* **254**, 4884 (1979); J. E. DeLarco *et al.*, *ibid.* **255**, 3685 (1980).
27. E. S. Kimball and T. C. Warren, *Biochim. Biophys. Acta* **771**, 82 (1984).

28. We thank L. Kissinger for help with the binding assays and P. Yoshida for assistance in manuscript preparation.

29 July 1988; accepted 18 November 1988