The collagenase gene family in humans consists of at least four members

Danièle MULLER,* Béatrice QUANTIN,† Marie-Claude GESNEL,† Régine MILLON-COLLARD,* Joseph ABECASSIS* and Richard BREATHNACH††

*Centre Régional de Lutte contre le Cancer Paul Strauss, 3 rue de la Porte de l'Hôpital, and

†Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS,

Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Faculté de Médecine, 11 rue Humann, 67085 Strasbourg Cédex, France

Stromelysin is a collagenase-related connective-tissue-degrading metalloproteinase. We have detected RNAs capable of hybridizing to a rat stromelysin cDNA in 11 of 69 human tumours tested. Molecular cloning of cDNAs to these RNAs has identified them as a mixture of stromelysin RNA and a transcript of a hitherto-undescribed related human gene, the stromelysin-2 gene. We have also isolated cDNAs corresponding to a more distantly related new human gene, the pump-1 gene. A comparison of the cDNA-derived amino acid sequences of stromelysin-2 and pump-1 with the known sequences of stromelysin and collagenase reveals significant similarities, with conservation of sequence motifs believed to have functional importance in metalloproteinase action. We conclude that the collagenase gene family in humans consists of at least four members, and speculate that expression of these genes plays a role in cancer.

INTRODUCTION

Tumour invasion and metastasis are of fundamental importance in cancer. It is a well established concept that these processes require an enzymic degradation of the host interstitial matrix. This concept is supported by many reports of increased proteolytic activities in tumour cells (reviewed in [1]). For instance, tumour cells release collagenases with specificity for interstitial collagens types I–III or basement-membrane type IV collagen (see [1–3] for references).

Stromelysin is a metalloproteinase related to collagenase (there is about 55 % similarity in their amino acid sequences) whose substrates include proteoglycans and fibronectin, but not type I collagen [4,5]. Little is known about the expression of stromelysin in human tumours. However, the apparent rat homologue of stromelysin, transin, is induced in rat-embryo fibroblasts by oncogenic transformation [6,7], and transin mRNA can be detected in mouse skin squamous-cell carcinomas, but not in premalignant benign papillomas or normal mouse epidermis [8]. These observations suggest that transin/stromelysin may be involved in tumour invasion in these rodent systems.

Given the crucial importance of tumour invasion and metastasis in cancer, and the role ascribed to proteinases in these phenomena, it was important to ascertain whether stromelysin RNA or related species could be detected in human tumours. In the present paper we report the results of a search for such species in 69 human tumours of various different types. This search has led to the identification of two new members of the human collagenase gene family.

METHODS

Tumour samples

RNA blotting was used to detect RNA species capable

of hybridizing to transin cDNA in various human tumours. Primary-tumour samples, 50 in all, were tested, including 32 adenocarcinomas (24 breast, three thyroid, two colon, one rectum, one lung and one endometrium), four upper-respiratory-tract carcinomas, two basal-cell carcinomas, one malignant melanoma, two sarcomas, one teratocarcinoma, four squamous-cell carcinomas of the lung, two squamous-cell carcinomas of the lip, one squamous-cell carcinoma of the forehead and one thyroid goitre. Samples were also prepared from 12 metastatic lymph nodes (derived from five breast, one laryngeal and one vulval carcinoma, three malignant melanomas and two Hodgkin's disease) and seven metastatic dermal nodules (derived from four breast carcinomas, two malignant melanomas and one liposarcoma), one omental metastasis of an endometrial adenocarcinoma and one nasal mucosa metastasis of a malignant melanoma. Normal tissues were obtained from macroscopically non-invaded lung, skin, colon, rectum and thyroid from nine subjects carrying carcinomas, the site of sampling being at least 4 cm away from the tumour mass.

Samples were quickly frozen and stored in liquid nitrogen after surgical excision. The histological type and grade of all tumour samples were examined on paraffin-embedded sections stained with haematoxylin and eosin.

RNA preparation and electrophoresis

Small pieces of frozen tissue were powdered with a Thermovac (Touzart et Matignon, Vitry sur Seine, France) and used for preparation of RNA by the LiCl/urea technique [9]. RNA samples (20 µg) were subjected to gel electrophoresis as described elsewhere [6,7] before transfer to Hybond-N nylon membranes (Amersham). Pre-hybridizations and hybridizations to nick-translated transin cDNA (the isolation of this probe has been

Abbreviations used: 1 × SSC, 0.15 M-NaCl/15 mM-sodium citrate; poly(A)⁺, polyadenylated.

† To whom correspondence should be addressed.

These sequence data have been submitted to the EMBL/GenBank Data Libraries under the accession number Y00728.

Fig. 1. Detection of a 1.9-kb RNA in human tumour RNA with a rat transin probe

Lanes 1 and 2, squamous-cell carcinomas of the larynx; lanes 3 and 4, thyroid carcinomas; lane 5, breast carcinoma; lane 6, subcutaneous metastasis of breast carcinoma; lane 7, nasal mucosa metastasis of a malignant melanoma; lane 8, squamous-cell carcinoma of the lung; lane 9, normal lung tissue; lane T contains RNA from polyoma-virus-transformed mouse fibroblasts [22].

described [6]) were carried out under conditions of moderate stringency. [$5 \times \text{SSC}/50\%$ (v/v) formamide] as described previously [6,7]. Membranes were washed in $0.36 \text{ M-NaCl}/20 \text{ mM-sodium phosphate}$, pH 7.7, containing 2 mM-EDTA and 0.1% SDS at 50°C . RNA was shown to be intact by re-hybridization of samples to a probe for a ubiquitous RNA [10].

Analysis of cloned cDNAs

Poly(A)⁺ RNA was prepared [11] from a pool of those RNA samples which contained material hybridizing to transin cDNA. The poly(A)⁺ RNA was used for the creation of a cDNA library in the vector $\lambda\text{gt}10$ [12], using the adaptor-tailing technique of Haymerle *et al.* [13]. About 10^6 recombinant phages were obtained from $4 \mu\text{g}$ of poly(A)⁺ RNA. Of these, 3×10^5 phages were screened for inserts capable of hybridizing to transin cDNA by standard techniques [11]. Hybridizations were carried out as for the RNA blotting experiments, except that low stringency conditions were used ($6 \times \text{SSC}/40\%$ formamide). Positively responding phages were plaque-purified and their cDNA inserts, bordered by restriction-*EcoRI* sites, transferred to the *EcoRI* site of plasmid pUN121 [14]. Resulting plasmids were amplified and used as a source of cloned cDNAs. Fragments of cloned cDNAs were transferred to the vectors M13tg130/131 [15] for sequencing by the dideoxy method [16] described previously. When restriction-enzyme fragments of the cDNAs were introduced into the M13 vectors, the universal sequencing primer was used for sequencing from the corresponding site. To fill some gaps, specific oligonucleotide sequencing primers were used whose sequences were based on those already obtained by using the universal primer. The sequencing strategies are shown in Figs. 2 and 3 (below). In this way all restriction sites used for sequencing were overlapped. The sequences of stromelysin-2 and pump-1 cDNAs shown in Figs. 2 and 3 (below) were determined on both strands.

RESULTS AND DISCUSSION

RNA blotting was used to identify species capable of hybridizing to a transin (rat stromelysin) cDNA probe in

RNA from various human tumours (see the Methods section for details). Six of the 50 primary tumours and five of the 19 metastases tested contained a 1.9-kb RNA detectable with the transin cDNA probe (see Fig. 1 for some examples). In primary tumours this RNA was found in six of 13 tumours tested with a squamous component, namely two upper-respiratory-tract carcinomas, one basal-cell carcinoma and squamous-cell carcinomas of the lung, lip and forehead (one of each). We did not detect the RNA in any of the 32 adenocarcinomas or two sarcomas, nor in any of the normal tissues tested. Among the five metastatic-tumour samples expressing the 1.9-kb RNA, one was derived from a primary tumour with a squamous component (squamous-cell carcinoma of the vulva), and the other three had metastasized to tissues with a squamous component (two dermal nodules of breast carcinomas and one metastasis of a malignant melanoma in the nasal cavity). The remaining positive sample was an omental metastasis of an endometrial adenocarcinoma.

Although transin appears to be the rat homologue of stromelysin, it was necessary to determine whether the 1.9-kb RNA was effectively stromelysin RNA. To this end we prepared a cDNA library in the bacteriophage vector $\lambda\text{gt}10$ from a mixture of RNAs from the 'positive' tumours and screened it by using the transin cDNA probe. Conditions of reduced stringency relative to those used to screen the RNA samples were chosen, in the hope of isolating cDNAs corresponding to all the stromelysin-related RNAs present in the tumour samples.

Screening of 3×10^5 plaques yielded 24 strongly positive plaques and nine weakly positive plaques. Analysis of the cDNA inserts of the positive phages by a combination of restriction-enzyme mapping, nucleotide sequencing and cross-hybridization experiments allowed us to identify four of the strongly positive cDNAs as stromelysin cDNA [5] and four of the weakly positive cDNAs as collagenase cDNAs [5,17]. The remaining 20 strongly positive cDNAs appeared to be derived from one new gene, which we decided to call the 'stromelysin-2' gene. The five non-collagenase weakly positive cDNAs appeared to be derived from a second new gene, which we decided to call the 'pump-1' (for putative metalloproteinase) gene. From the frequency of stromelysin versus stromelysin-2 cDNA clones obtained we conclude that the 1.9-kb RNAs seen in Fig. 1 are predominantly stromelysin-2 RNAs. It should be stressed that the higher stringency conditions of hybridization used for the RNA blotting experiments (compared with the phage screening experiments) would not have allowed detection of pump-1 or collagenase RNAs, which are thus absent from Fig. 1. The pump-1 cDNA probe will detect a 1.2-kb RNA in the pool of RNAs used to create the cDNA library, however (results not shown).

The longest stromelysin-2 (1.7 kb) and pump-1 (1.1 kb) cDNAs appear to be complete at their 3'-ends, as they carry a poly(A) tail (results not shown) preceded by the polyadenylation signal AATAAA [18]. The stromelysin-2 cDNA has a coding capacity for a putative protein of 476 amino acids, which we term 'stromelysin-2' (Fig. 2). The pump-1 cDNA has coding capacity for a putative protein of 267 amino acids, which we term 'pump-1' (Fig. 3). The sequences of these two proteins are compared in Fig. 4 with those of collagenase and stromelysin. This comparison reveals that stromelysin-2 is similar, over its entire length, to these enzymes (overall

22
5'...AAAGAAGGTAAGGGCAGTGAGA

23
ATG ATG CAT CTT GCA TTC CTT GTG CTG TTG TGT CTG CCA GTC TGC TCT GCC TAT CCT CTG AGT GGG GCA GCA AAA GAG GAG GAC TCC AAC
MET MET HIS LEU ALA PHE LEU VAL LEU LEU CYS LEU PRO VAL CYS SER ALA TYR PRO LEU SER GLY ALA ALA LYS GLU GLU ASP SER ASN

113
AAG GAT CTT GCC CAG CAA TAC CTA GAA AAG TAC TAC AAC CTC GAA AAG GAT GTG AAA CAG TTT AGA AGA AAG GAC AGT AAT CTC ATT GTT
LYS ASP LEU ALA GLN GLN TYR LEU GLU LYS TYR TYR ASN LEU GLU LYS ASP VAL LYS GLN PHE ARG ARG LYS ASP SER ASN LEU ILE VAL

203
AAA AAA ATC CAA GGA ATG CAG AAG TTC CTT GGG TTG GAG GTG ACA GGG AAG CTA GAC ACT GAC ACT CTG GAG GTG ATG CGC AAG CCC AGG
LYS LYS ILE GLN GLY MET GLN LYS PHE LEU GLY LEU GLU VAL THR GLY LYS LEU ASP THR ASP THR LEU GLU VAL MET ARG LYS PRO ARG

293
TGT GGA GTT CCT GAC GTT GGT CAC TTC AGC TCC TTT CCT GGC ATG CCG AAG TGG AGG AAA ACC CAC CTT ACA TAC AGG ATT GTG AAT TAT
CYS GLY VAL PRO ASP VAL GLY HIS PHE SER SER PHE PRO GLY MET PRO LYS TRP ARG LYS THR HIS LEU THR TYR ARG ILE VAL ASN TYR

383
ACA GAT TTG CCA AGA GAT GCT GTT GAT TCT GCC ATT GAG AAA GCT CTG AAA GTC TGG GAA GAG GTG ACT CCA CTC ACA TTC TCC AGG
THR PRO ASP LEU PRO ARG ASP ALA VAL ASP SER ALA ILE GLU LYS ALA LEU LYS VAL TRP GLU GLU VAL THR PRO LEU THR PHE SER ARG

473
CTG TAT GAA GGA GAG GCT GAT ATA ATG ATC TCT TTC GCA GTT AAA GAA CAT GGA GAC TTT TAC TCT TTT GAT GGC CCA GGA CAC AGT TTG
LEU TYR GLU GLY GLU ALA ASP ILE MET ILE SER PHE ALA VAL LYS GLU HIS GGA ASP PHE TYR SER PHE ASP GLY PRO GLY HIS SER LEU

563
GCT CAT GCC TAC CCA CCT GGA CCT GGG CTT TAT GGA GAT ATT CAC TTT GAT GAT GAT GAA AAA TGG ACA GAA GAT GCA TCA GGC ACC AAT
ALA HIS ALA TYR PRO PRO GLY PRO GLY LEU TYR GLY LEU TYR GLY LEU PHE HIS SER ALA ASN THR GLU ALA LEU MET TYR PRO LEU TYR ASN

653
TTA TTC CTC GTT GCT GCT CAT GAA CTT GGC CAC TCC CTG GGG CTC TTT CAC TCA GCC AAC ACT GAA GCT TTG ATG TAC CCA CTC TAC AAC
LEU PHE LEU VAL ALA ALA HIS GLU LEU GLY HIS SER LEU GLY LEU PHE HIS SER ALA ASN THR GLU ALA LEU MET TYR PRO LEU TYR ASN

743
TCA TTC ACA GAG CTC GCC CAG TTC CGC CTT TCG CAA GAT GAT GTG AAT GGC ATT CAG TCT CTC TAC GGA CCT CCC CCT GCC TCT ACT GAG
SER PHE THR GLU LEU ALA GLN PHE ARG LEU SER GLN ASP ASP VAL ASN GLY ILE GLN SER LEU TYR GLY PRO PRO PRO ALA SER THR GLU

833
GAA CCC CTG GTG CCC ACA AAA TCT GTT CCT TCG GGA TCT GAG ATG CCA GCC AAG TGT GAT CCT GCT TTG TCC TTC GAT GCC ATC AGC ACT
GLU PRO LEU VAL PRO THR LYS SER VAL PRO SER GLY SER GLU MET PRO ALA LYS CYS ASP PRO ALA LEU SER PHE ASP ALA ILE SER THR

923
CTG AGG GGA GAA TAT CTG TTC TTT AAA GAC AGA TAT TTT TGG CGA AGA TCC CAC TGG AAC CCT GAA CCT GAA TTT CAT TTG ATT TCT GCA
LEU ARG GLY GLU TYR LEU PHE PHE LYS ASP ARG TYR PHE TRP ARG ARG SER HIS TRP ASN PRO GLU PRO GLU PHE HIS LEU ILE SER ALA

1013
TTT TGG CCC TCT CTT CCA TCA TAT TTG GAT GCT GCA TAT GAA GTT AAC AGC AGG GAC ACC GTT TTT ATT TTT AAA GGA AAT GAG TTC TGG
PHE TRP PRO SER LEU PRO SER TYR LEU ASP ALA ALA TYR GLU VAL ASN SER ARG ASP THR VAL PHE ILE PHE LYS GLY ASN GLU PHE TRP

1103
GCC ATC AGA GGA AAT GAG GTA CAA GCA GGT TAT CCA AGA GGC ATC CAT ACC CTG GGT TTT CCT CCA ACC ATA AGG AAA ATT GAT GCA GCT
ALA ILE ARG GLY ASN GLU VAL GLN ALA GLY TYR PRO ARG GLY ILE HIS THR LEU GLY PHE PRO PRO THR ILE ARG LYS ILE ASP ALA ALA

1193
GTT TCT GAC AAG GAA AAG AAG AAA ACA TAC TTC TTT GCA GCG GAC AAA TAC TGG AGA TTT GAT GAA AAT AGC CAG TCC ATG GAG CAA GGC
VAL SER ASP LYS GLU LYS LYS LYS THR TYR PHE PHE ALA ALA ASP LYS TYR TRP ARG PHE ASP GLU ASN SER GLN SER MET GLU GLN GLY

1283
TTC CCT AGA CTA ATA GCT GAT GAC TTT CCA GGA GTT GAG CCT AAG GTT GAT GCT GTA TTA CAG GCA TTT GGA TTT TTC TAC TTC TTC AGT
PHE PRO ARG LEU ILE ALA ASP ASP PHE PRO GLY VAL GLU PRO LYS VAL ASP ALA VAL LEU GLN ALA PHE GLY PHE PHE TYR PHE PHE SER

1373
GGA TCA TCA CAG TTT GAG TTT GAC CCC AAT GCC AGG ATG GTG ACA CAC ATA TTA AAG AGT AAC AGC TGG TTA CAT TGC TAG GCG AGA TAG
GLY SER SER GLN PHE GLU PHE ASP PRO ASN ALA ARG MET VAL THR HIS ILE LEU LYS SER ASN SER TRP LEU HIS CYS ***

Fig. 2. Partial nucleotide sequence and predicted amino acid sequence of stromelysin-2 cDNA

The first 1462 nucleotides of the 1.7-kb cDNA are shown. A restriction-endonuclease map of the cDNA is shown below the sequence. Arrows indicate the extent and direction of sequences determined using either specific primers (■→) or the universal primer (other arrows). Abbreviations for restriction endonucleases: S, *ScaI*; B, *SphI*; N, *NsiI*; H, *HindIII*; T, *SstI*; D, *DraI*; A, *HpaI*; P, *PvuII*; V, *EcoRV*; R, *EcoRI*.

similarity of 52 and 78 % to collagenase and stromelysin respectively), and this strongly suggests that the complete sequence of stromelysin-2 was present on the stromelysin-2 cDNA sequenced. Consistent with this hypothesis, stromelysin-2 has a hydrophobic *N*-terminal sequence as a candidate signal sequence.

Pump-1 is significantly shorter (267 amino acids) than stromelysin or collagenase (477 and 469 amino acids respectively). Judging from the alignment shown in Fig. 4, the 'missing' sequences appear to be C-terminal. Like stromelysin-2, pump-1 has a candidate signal sequence, suggesting that the *N*-terminal sequence may be complete.

Fig. 3. Partial nucleotide sequence and predicted amino acid sequence of pump-1 cDNA

The first 837 nucleotides of the 1.1-kb cDNA are shown. A restriction-endonuclease (*Xho*I, *Sst*I and *Xmn*I) map of the cDNA is shown below the sequence. Symbols are as in the legend to Fig. 2.

The 'missing' C-terminal sequences are not due to a cloning artefact, as the same sequence was determined from two independent cDNA clones, both of which contained a stop codon after the codon coding for amino acid 267 of pump-1, and an equivalent 3'-untranslated region. A comparison of the sequence of pump-1 with the corresponding regions of stromelysin and collagenase reveals a sequence similarity of 49 and 44 % respectively. Stromelysin-2 appears thus to be a normal, and pump-1 a truncated, member of the collagenase family.

Although a direct experimental test is necessary to determine whether pump-1 and stromelysin-2 are secreted proteinases, their sequences are consistent with this possibility. As mentioned above, both proteins have candidate signal sequences. In addition, both proteins contain the region corresponding to the putative zinc-chelating region of stromelysin and collagenase, with conservation of the histidine residues (* in Fig. 4) believed to participate in binding the zinc atom known to be at the active site of all metalloproteinases studied in any detail [5].

Stromelysin and collagenase are secreted as pro-enzymes [5,19]. The amino acids found between the signal

sequence and the phenylalanine at the N-termini of the activated forms of these enzymes (see Δ , Fig. 4) are cleaved off during activation. Just before these phenylalanine residues lies a conserved block of eight amino acids (underlined in Fig. 4). These amino acids are strictly conserved in stromelysin-2 and pump-1. Stromelysin-2 and pump-1 have thus all the hallmarks of secreted activable metalloproteinases of the collagenase family.

The 'truncated' structure of pump-1 is of particular interest in view of the fact that transin (rat stromelysin) comprises two separate domains. We showed previously that the C-terminal moiety of transin (amino acids 288-475) has weak, but statistically significant, similarity to hemopexin [7]. This observation holds true also for collagenase and stromelysin (results not shown, but the beginning of this moiety of these proteins is shown by an arrow marked 'C' in Fig. 4). We have subsequently shown that the N-terminal moiety of transin (amino acids 1-273; see the legend to Fig. 4 and the arrow marked N) houses its proteinase activity (R. Breathnach, unpublished work). The 267 amino acids of pump-1 correspond to those present in the N-terminal

Fig. 4. Comparison of the predicted amino acid sequences of pump-1, stromelysin, stromelysin-2 and collagenase

Amino acids conserved in all the sequences shown are marked by colons (:). ▲ Denotes the signal-peptidase cleavage site of procollagenase [5,17]. △ Indicates the *N*-terminus of activated collagenase [5,19]. A conserved sequence of eight amino acids just upstream is underlined. The arrow marked 'N' shows where the end of the *N*-terminal moiety of transin (rat stromelysin) lines up on the stromelysin and collagenase sequences. This moiety can act as a proteinase. The arrow marked 'C' shows the beginning of the *C*-terminal hemopexin-like moiety of collagenase family enzymes.

moiety of transin, i.e. pump-1 has a potential proteinase domain but not a hemopexin-like domain. The function of the hemopexin-like domain of collagenase-family proteinases is unknown. We may speculate that it might be involved in controlling enzyme activation, conferring substrate selectivity or binding of inhibitors. The absence of this domain from pump-1 would in these cases have important implications.

In conclusion, we have described two new members of the collagenase gene family, which thus consists of at least four members in humans: genes for collagenase, stromelysin, stromelysin-2 and pump-1. The parent family member, collagenase, is an important enzyme which plays a role in various pathological conditions such as cancer and rheumatoid arthritis [1]. The discovery of new collagenase family members is thus of some interest, and it will be of importance to characterize the proteins stromelysin-2 and pump-1 and to investigate the factors controlling their expression.

Rat stromelysin RNA is induced by oncogenes in cultured cells [6,7]. Many human tumours contain activated oncogenes [20]. It is thus of interest that transcripts of the stromelysin or stromelysin-2 genes can be detected in some human tumours. These transcripts were preferentially detected in tumours with a squamous component or in metastases of tumours in epidermal tissues. Tumour samples are heterogeneous, and mesenchymal cells (fibroblasts, macrophages) are often found in high amounts in the vicinity of tumour cells. Production and release of proteinases could thus result from tumour-cell-host-tissue interactions [21]. The

material described here should allow use of '*in situ*' hybridization techniques to address this question. It is interesting to note that most of the stromelysin/stromelysin-2-positive tumours are rapidly evolutive (head and neck, spinal-cell carcinoma) or advanced-stage diseases (poorly differentiated lung squamous-cell carcinomas). These preliminary results raise the question as to whether expression of these genes may contribute to the highly malignant state of epidermoid carcinomas. Further analysis of a wide spectrum of these tumours and extension of these studies to expression of the pump-1 gene is called for.

We thank Mr. F. Ruffenach and Mr. A. Staub for synthesis of oligonucleotide primers for sequencing. This work was supported by grants from the CNRS, INSERM, the Association pour la Recherche contre le Cancer n° 6305 and the Ligue Nationale Française contre le Cancer (Fédération Nationale and Comité Départemental du Bas-Rhin).

REFERENCES

1. Mullins, D. E. & Rohrich, S. T. (1983) *Biochim. Biophys. Acta* **695**, 177-214
2. Wooley, D. E. (1984) *Cancer Metastasis Review* **3**, 361-372
3. Liotta, L. A. (1986) *Cancer Res.* **46**, 1-7
4. Chin, J. R., Murphy, G. & Werb, Z. (1985) *J. Biol. Chem.* **260**, 12367-12376
5. Whitham, S. E., Murphy, G., Angel, P., Rahmsdorf, H.-J., Smith, B. J., Lyons, A., Harris, T. J., Reynolds, J. J., Herrlich, P. & Docherty, A. J. P. (1986) *Biochem. J.* **240**, 913-916

6. Matrisian, L. M., Glaichenhaus, N., Gesnel, M.-C. & Breathnach, R. (1985) *EMBO J.* **4**, 1435-1440
7. Matrisian, L. M., Leroy, P., Ruhlmann, C., Gesnel, M.-C. & Breathnach, R. (1986) *Mol. Cell. Biol.* **6**, 1679-1686
8. Matrisian, L. M., Bowden, G. T., Krieg, P., Furstenberger, G., Briand, J.-P., Leroy, P. & Breathnach, R. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **38**, 9413-9417
9. Auffray, C. & Rougeon, F. (1980) *Eur. J. Biochem.* **107**, 303-314
10. Masiakowski, P., Breathnach, R., Bloch, J., Gannon, F., Krust, A. & Chambon, P. (1982) *Nucleic Acids Res.* **10**, 7895-7903
11. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
12. Huynh, T. V., Young, R. A. & Davis, R. W. (1985) in *DNA Cloning Techniques* (Glover, D. M., ed.), vol. 1, pp. 49-78, IRC Press, Oxford
13. Haymerle, H., Herz, J., Bressan, G. M., Frank, R. & Stanley, K. K. (1986) *Nucleic Acids Res.* **14**, 8615-8624
14. Nilsson, B., Uhlen, M., Josephson, S., Gatenbeck, S. & Philipson, L. (1983) *Nucleic Acids Res.* **11**, 8019-8030
15. Kieny, M. P., Lathe, R. & Lecoq, J. P. (1983) *Gene* **26**, 91-99
16. Sanger, F., Nicklen, S. & Coulson, A. R. (1977). *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463-5467
17. Goldberg, G. I., Wilhelm, S. M., Kronberger, A. M., Bauer, E. A., Grant, G. A. & Eisen, A. Z. (1986) *J. Biol. Chem.* **261**, 6600-6605
18. Proudfoot, N. & Brownlee, G. G. (1981) *Nature (London)* **252**, 359-362
19. Grant, G. A., Eisen, A. Z., Marmer, B. L., Roswit, W. T. & Goldberg, G. I. (1987) *J. Biol. Chem.* **262**, 5886-5889
20. Bishop, J. M. (1987) *Science* **235**, 305-311
21. Biswas, C. (1984) *Cancer Lett.* **24**, 201-207
22. Rautmann, G., Glaichenhaus, N., Nahgashfar, Z., Breathnach, R. & Rassoulzadegan, M. (1982) *Virology* **122**, 306-317

Received 19 October 1987/9 February 1988; accepted 18 March 1988